

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 569)  
 AUTHORS NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
 JOURNAL Tumor Gene Index  
 COMMENT Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. cDNA library Preparation: M.B. Soares Lab Clone distribution:  
 NCI-CCGAP clone distribution information can be found through the  
 I.M.A.G.E. Consortium/LLNL at:  
 www-bio.llnl.gov/bbrp/image/image.html The following repetitive  
 elements were found in this cDNA sequence: 1-22,  
 >Alu-rich/Low-complexity  
 Seq primer: M13 Forward  
 POLYA=yes.

## FEATURES

source  
 1..569  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3087997"  
 /clone\_lib="NCI-CCGAP\_Subs"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker. Site 1: Not 1; Site 2: Eco RI; NCI-CCGAP Sub5  
 is a subtracted library derived from NCI-CCGAP\_Subs5. The  
 NCI-CCGAP\_Subs8 library had 2.5 million recombinants. A  
 single-stranded DNA preparation of NCI-CCGAP\_Subs5 was used  
 as a tracer in a subtractive hybridization with a driver  
 clone library derived from NCI-CCGAP\_Subs5 (IMAGE  
 clone Ids 2732833-2737415, 3068040-3069191; 25% of the  
 driver population). A pool of clones from NCI-CCGAP\_Subs4  
 (IMAGE clone Ids 2723592-2729326; 25% of the driver  
 population), NCI-CCGAP\_Subs5 (pool AIF-AJU, IMAGE Ids  
 2728999-2733190; 25% of the driver population), and  
 NCI-CCGAP\_Subs7 (IMAGE Ids 3069192-3072238, 3081854-3084550  
 : 25% of the driver population). Subtraction was  
 performed as previously described [Bonaldo, Lennon &  
 Soares (1996)]: Normalization and Subtraction: Two  
 Approaches To Facilitate Gene Discovery. Genome Research  
 6, 791-806.  
 TAG\_LIB=NCI-CCGAP\_GC4  
 TAG\_TISSUE=germ cell  
 TAG\_SEQ=AAATC"

BASE COUNT 149 a 142 c 95 g 183 t  
 ORIGIN

Query Match 85.2% Score 397; DB 10; Length 569;  
 Fast Local Similarity 98.0%; Pred. No. 3.2e-105;  
 Matches 444; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

OY 1 atcaagaacataagattcggacaataatacttcctaccatccaccccaaatctta 60  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 95 ATCAAGAACAATAGAGTII-GGGCAATATACITTCATCCATCCACCCCAATCTTA 153  
 OY 61 ctctactcattctctctcattatttgggaataatcatcagaagatgtgtcttaagta 120  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 154 CTTACTATCTCATCTCATTAATTTGGGAATATCATGAGAGATGTGTCTTGAGTA 213  
 OY 121 agagattaaagaataaagcttttttgaacccctcccaaccccccatgccagggtgtcac 180  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 214 AGAGATTAAAGAAATATAGCTTTTTCACCCCTGCCAACACCCCATCCCGGGTGTAC 273  
 OY 181 cctccatacaataacatccagagaagaataagtcccttctctatccgtaattctgcc 240  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 274 CTTCCAAATACATAGATGCCGGAAGAGTAGTTCCTCTTCTGTATGCCGTAATCTGCC 333

OY 241 atcatcttccatcttccagttctctcttccattgcaagtccacaaatctggtctcaggat 300  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 334 ATCATCTTCCATCTTCCAGTCT-CTTTCATTCGAAGTCACATCTGGGTCTCAGGAT 392  
 OY 391 tataccgctattatctgcatcttcttcttcttcttcttcttcttcttcttcttctt 360  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 393 TATACCGCTTATAGTCTCGATCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 452  
 OY 361 ctggagagagtcctcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 420  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 453 CT-GGAGAGGTGCTCTTTCCTTCATCCTGACCTGACCTGACCTGACCTGACCTGAC 511  
 OY 421 catcatcgggtcttccacaccttcttcttcttcttcttcttcttcttcttcttctt 453  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 512 TATCACTTGGGCTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 543  
 RESULT  
 BF511148  
 LOCUS  
 DEFINITION  
 BF511148  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 human.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 571)  
 AUTHORS NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
 JOURNAL Tumor Gene Index  
 COMMENT Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. cDNA library Preparation: M.B. Soares Lab Clone distribution:  
 NCI-CCGAP clone distribution information can be found through the  
 I.M.A.G.E. Consortium/LLNL at:  
 www-bio.llnl.gov/bbrp/image/image.html The following repetitive  
 elements were found in this cDNA sequence: 1-22,  
 >Alu-rich/Low-complexity  
 Seq primer: M13 Forward  
 POLYA=yes.

## FEATURES

source  
 1..571  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3087989"  
 /clone\_lib="NCI-CCGAP\_Subs8"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker. Site 1: Not 1; Site 2: Eco RI; NCI-CCGAP\_Subs5  
 is a subtracted library derived from NCI-CCGAP\_Subs5. The  
 NCI-CCGAP\_Subs8 library had 2.5 million recombinants. A  
 single-stranded DNA preparation of NCI-CCGAP\_Subs5 was used  
 as a tracer in a subtractive hybridization with a driver  
 clone library derived from NCI-CCGAP\_Subs5 (IMAGE  
 clone Ids 2732833-2737415, 3068040-3069191; 25% of the  
 driver population). A pool of clones from NCI-CCGAP\_Subs1  
 (IMAGE clone Ids 2723592-2729326; 25% of the driver  
 population), NCI-CCGAP\_Subs6 (pool AIF-AJU, IMAGE Ids  
 2728999-2733190; 25% of the driver population), and  
 NCI-CCGAP\_Subs7 (IMAGE Ids 3069192-3072238, 3081854-3084550  
 : 25% of the driver population). Subtraction was  
 performed as previously described [Bonaldo, Lennon &  
 Soares (1996)]: Normalization and Subtraction: Two  
 Approaches To Facilitate Gene Discovery. Genome Research  
 6, 791-806.  
 TAG\_LIB=NCI-CCGAP\_GC4

[illegible]

```

Db 441 CT-CGAGGAGTGCTCTCINTGCTCTATCACCCTGACTCTCCACAAGAACCAAGGGCAGCTCCT 499

Q: 421 catcctggggttcaccactttcaggggtaag 453
||||| ||||||||| ||||| |||
Db 500 CATCAGT-GGCTTACCACATTCAGGGTAAG 531

RESULT 7
A1656143 A1656143 750 bp mRNA linear EST 04-MAY-1999
DEFINITION t143f03.x1 NCI_CGAP.GC6 Homo sapiens cDNA clone IMAGE:2243549 3'
similar to TR:O15205 O15205 DIUBIQUITIN.1, mRNA sequence.
ACCESSION A1656143
VERSION A1656143.1 GI:4740122
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 750)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute. Cancer genome Anatomy Project (CGAP).
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: coapbs-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
***bio.llnl.gov/hbrp/image/image.html
Seq primer: -400p from Gibco
High quality sequence stop: 450.
Location/Qualifiers
1..750
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2243549"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker: Site.1: Not I; Site.2: Eco RI; Plasmid DNA
from the normalized library NCI_CGAP_GC4 was prepared. and
ss circles were made in vitro. following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneIDs
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 197 a 188 c 142 g 213 t 10 others
ORIGIN
Query Match 85.0% Score 396; DB 9; Length 750;
Best local Similarity 97.8% Pred. No. 6.Be-105;
Matches 443; Conservative 0; Mismatches 6; Indels 4; Gaps 4;
Q: 1 atcaagaacatagagttcgggccaatactctcatctaccatccaccaccaatctta 60
||||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 83 A1CAAAGAACATAGAGTT-GGGCAATATATCTTCATCTACCCATCCACCCAATCTTA 141
||||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Q: 61 ctctactcatctcattctcattgaatttgggaatcatcagaagatgtctcgttaagta 120
||||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 142 C1CTACTATCTCATCTCATTAATTTTGGGAATCATCAGAGATGTGTTCTGTAGTA 201
||||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Q: 101 agagattaaagaataaagctttttgacccctgcacaaccccccatgccaggatgtcac 180
||||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 202 AGAGATTAAAGAATAAAGCTTTTIGACCTCGCAACACCCATCCCAAGGGTGGTTCAC 261
||||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 181 cctccataataacatgcccaggaagagtaagttgccccttctgcatcccatcccaaaattta 240
DB 262 cctccaaatacaaaataagatgcccaggaagagtaagttgccccttctgcatcccatcccaaaattta 321
QY 241 atcattcttcccatcttcagagtccttcccatgcaaatcacaatctgggtctcaggat 300
DB 322 atcattcttcccatcttcagagtccttcccatgcaaatcacaatctgggtctcaggat 380
QY 301 tatcccgctttagtctgagatcattgcttccatctgctccactgagctggagccttcgac 360
DB 381 tatcccgctttagtctgagatcattgcttccatctgctccactgagctggagccttcgac 440
QY 361 ctggagaagatcacctcttgcctcaccctgacctgacctcacaagaacaaagagcagctct 420
DB 441 ct-ggaggaggtgccccttctgctcaccctgacctcacaagaacaaagagcagctct 499
QY 421 catcactggggttcaccacttccagggtaag 453
DB 500 catcact-ggggttcaccactcctcagggtaag 531

RESULT 8
AA583465 511 bp mRNA linear EST 06-SEP-1997
LOCUS nt39d07.s1 NCI_CGAP_G05 Homo sapiens cDNA clone IMAGE:1086157 3'
DEFINITION similar to TR:G972017 G972017 mRNA; EXPRESSED SEQUENCE 123 1; mRNA
sequence.
ACCESSION AA583465.1 GI:2368074
VERSION AA583465
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 511)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 381.
FEATURES
Location/Qualifiers
1..511
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1086157"
/clone_lib="NCI_CGAP_G05"
/tissue_type="germ cell tumor"
/lab_host="SOLR (kanamycin resistant)"
/notes="Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XbaI
; Cloned unidirectionally. Primer: oligo dT. Mixed germ
cell tumors. 5' adaptor sequence: 5' GAATTCGGACGAG 3' 2.
adaptor sequence: 5' CTCGAGTTTTTTTTTTTT 3' Average
insert size: 0.7 kb."
BASE COUNT 141 a 133 c 84 g 153 t
ORIGIN

```

Query Match 84.5%; Score 384; DB 9; Length 511;  
Best Local Similarity 99.3%; Pred. No. 2,3e-104;  
Matches 427; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

```

QY 1 atcaaaagaacatagaattcgggcaataatacttcatctaccatcccaaaattta 60
DB 79 atcaaaagaacatagaattt-gggcaataatacttcatctaccatcccaaaatttita 137
QY 61 ctctactcatctctctcatttaattttgggaaatcatcagaagatgttctggttgagta 120
DB 138 ctctactcatctctcatttcatatttttgggaaatcatcagaagatgttctggttgagta 197
QY 121 agagattaaagaataagctttttgaccccttgcacacacccccatgccaggggtggtcac 180
DB 192 agagattaaagaataaagctttttgaccccttgcacacacccccatgccaggggtggtcac 257
QY 191 ccttcaat'acaataacatgccaggaagagtaagttgccccttctgcatcccatcccaaaattta 240
DB 259 ccttcaat'acaataacatgccaggaagagtaagttgccccttctgcatcccatcccaaaattta 317
QY 241 atcattcttcccatcttcagagtccttccatctgcaatctcagagtcaggtctcaggat 300
DB 319 atcattcttcccatcttcagagtccttccatctgcaatctcagagtcaggtctcaggat 375
QY 301 tatcccgctttagtctgagatcattgcttccatctgctccactgagctggagccttcgac 360
DB 377 tatcccgctttagtctgagatcattgcttccatctgctccactgagctggagccttcgac 435
QY 361 ctggagaagatgaccttcttgcctcaccctgacctcacaagaacaaagagcagctct 420
DB 437 ct-ggaggaggtgccccttctgctcaccctgacctcacaagaacaaagagcagctct 495
QY 421 catcactggggttcaccacttccagggtaag 453
DB 496 catcactggggttcaccacttccagggtaag 505

RESULT 9
AA583447 568 bp mRNA linear EST 06-DEC-2000
LOCUS IMAGE:2087987 3' mRNA sequence.
DEFINITION UT-H-B14-aph-h-02-0-UI.s1 NCI_CGAP_Sub8 Homo sapiens cDNA clone
IMAGE:2087987 3', mRNA sequence.
ACCESSION BF511147
VERSION BF511147
KEYWORDS EST.
SOURCE EST.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 568)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
NCI-CGAP clone distribution information can be found through the
I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html The following repetitive
elements were found in this cDNA sequence: 1-22.
>Al_richlow_complexity
Seq primer: M13 Forward
POLYA=yes.
Location/Qualifiers
1..568
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3087987"
/clone_lib="NCI_CGAP_Sub8"
/lab_host="DH10B (Life technologies)"
/notes="Vector: p713D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; NCI_CGAP_Sub8
is a subtracted library derived from NCI_CGAP_Sub5. The

```

NCI\_CGAP\_Sub8 library had 2.5 million recombinants. A single-stranded DNA preparation of NCI\_CGAP\_Sub5 was used as a tracer in a subtractive hybridization with a driver comprising a pool of clones from NCI\_CGAP\_Sub5 (IMAGE clone IDs 2732833-2737415, 3058040-3059211; 25% of the driver population), a pool of clones from NCI\_CGAP\_Sub4 (IMAGE clone IDs 2723592-2729326; 25% of the driver population), NCI\_CGAP\_Sub5 (pool AIF-AJL; IMAGE IDs 2728959-2733190; 25% of the driver population), and NCI\_CGAP\_Sub7 (IMAGE IDs 3069192-3072328, 3081864-3084550; 25% of the driver population). Subtraction was performed as previously described [Bonaiuto, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806].

TAG\_LIB=NCI\_CGAP\_GC4  
TAG\_TISSUE=germ cell  
TAG\_SEQ=AAATC

BASE COUNT 150 a 141 c 95 g 182 t  
ORIGIN

Query Match 84.5%; Score 393.8; DB 10; Length 568;  
Best Local Similarity 97.6%; Pred. No. 2.7e-104;  
Matches 442; Conservative 0; Mismatches 7; Indels 4; Gaps 4;  
QY 1 atcaagaacacatagattcggcaatactcttcacccatccacacgaaatctta 60  
Db 95 ATCAAGAACAATAGAGTT-GGGCAATATATCTTACCTACCCATCCACCCAAATCTTA 153  
QY 61 ctctact 120  
Db 154 CTCTACTCATCTCATCTCATCTTAATTTTGGGAATCATCAGAAGATGTGTCTTGAGTA 213  
QY 121 agagattaaagaataagctttttgacccctgccacaccccatccacagagatgctgcac 180  
Db 214 AGAGATTAAGAATAATAGCTTTTTCACCCCTGCCAACACCCCATCCCGAGGTGGTAC 273  
QY 181 ctcccaatacaatacaccaggaagaatgcttccctctctctctctctctctctctctctc 240  
Db 274 CTCCCAATACAAATAGATGCCAGGAAGATGAAGTTGCCCTTTCTGATGCCGTAACTGCC 333  
QY 241 atcatct 300  
Db 334 ATCATCTTCCCAATCTCCAGTCT-CTTTCCATTCGAATCACAATCTGGGTCTCAGGAT 362  
QY 301 tataccgtcttagctcgtcattgcttccactgtgcccactgagctggagcttgcacac 360  
Db 393 TATACCCGCTTAGTCTCGATCATGTCTTACCTTGCCACTGAGCTTGGACCTTGGCAC 452  
QY 361 ctggagagagtgctctcttgcctctacacctgactccacagaacaaagggagctctct 420  
Db 453 CT-GGAGGAGGTGCTCTTTGGCTCATCTACCTGACTCCACAAGAACAAGGTGGCTGCT 513  
QY 421 catcactggagcttccacacttccagggttaag 453  
Db 512 CATCACT-GGGCTTCACCACTTTCAAGGTAAAG 543

RESULT 10

AW474876

LOCUS

DEFINITION AW474876 593 bp mRNA linear EST 24-FEB-2000  
similar to IR:015205 O15205 DIUBIOUTIN. 1; mRNA sequence.

ACCESSION

AW474876

VERSION

AW474876.1

KEYWORDS

EST

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 593)

AUTHORS

NCI\_CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE

Journal Tumor Gene Index

COMMENT

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: [cgapsb-ref@mail.nih.gov](mailto:cgapsb-ref@mail.nih.gov)

Life Technologies catalog #: 11547-015

DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

Seq primer: -400p from Gibco

High quality sequence stop: 416.

FEATURES

Source

1..593

Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:2852514"

/clone\_lib="NCI\_CGAP\_Lym12"

/tissue\_type="lymphoma, follicular mixed small and large cell"

/lab\_host="DH10B"

/note="Organ: lymph node; Vector: pCMV-SPORT6; Site:1;

Sal1; Site\_2: NotI; Cloned unidirectionally. primer:

Oligo dT. Average insert size 1.25 kb. Life Technologies

catalog #: 11547-015"

BASE COUNT 157 a 150 c 105 g 181 t

ORIGIN

Query Match 83.8%; Score 390.6; DB 9; Length 593;  
Best Local Similarity 97.1%; Pred. No. 2.4e-103;  
Matches 440; Conservative 0; Mismatches 9; Indels 4; Gaps 4;  
QY 1 atcaagaacacatagattcggcaatactcttcacccatccacacgaaatctta 60  
Db 90 ATCAAGAACAATAGAGTT-GGGCAATATATCTTACCTACCCATCCCAATCTTA 148  
QY 61 ctctact 120  
Db 143 CTCTACTCATCTCATCTCATTAATTTTGGGAATCATCAGAAGATGTGTCTTGAGTA 208  
QY 121 agagattaaagaataagctttttgacccctgccacaccccatccacagagtgctgcac 180  
Db 213 AGAGATTAAGAATAAACCTTTTTCACCCCTGCCAACACCCCATCCCGAGGTGGTAC 258  
QY 181 ctcccaatacaatacaccaggaagaatgcttccctctctctctctctctctctctctc 240  
Db 249 CTCCCAATACAAATAGATGCCAGGAAGATGAAGTTGCCCTTTCTGATGCCGTAACTGCC 328  
QY 241 atcatc 300  
Db 329 ATCATCTTCCCAATCTCCAGTCT-CTTTCCATTCGAATCACAATCTGGGTCTCAGGAT 387  
QY 301 tataccgtcttagctcgtcattgcttccactgtgcccactgagctggagcttgcacac 360  
Db 393 TATACCCGCTTAGTCTCGATCATGTCTTACCTTGCCACTGAGCTTGGACCTTGGCAC 447  
QY 361 ctggagagagtgctcttgcctctacacctgactccacagaacaaagggagctctct 420  
Db 449 CT-GGAGGAGGTGCTCTTTGGCTCATCTACCTGACTCCACAAGAACAAGGAGCTCT 506  
QY 421 catcactggagcttccacacttccagggttaag 453  
Db 507 CATCACT-GGGCTTCACCACTTTCAAGGTAAAG 538

RESULT 11

AV699595

LOCUS

DEFINITION

AV699595

AV699595

AV699595

AV699595

AV699595

AV699595

AV699595

AV699595

AV699595

AV699595

```

SOURCE
ORGANISM      human.
Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
1 (bases 1 to 755)
Xiao.H., Qiu.J., Liu.F., Huang.Q., Cheng.Z., Li.N., Du.J., Hu.W.,
Shen.K., Lu.G., Fu.G., Zhong.M., Xu.S., Gu.W., Huang.W., Zhao.X.,
Hu.G., Gu.J., Chen.Z., and Han.Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26): 15089-15094 (2001)
21525106
Contact: Zequang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@hgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
source
1..755
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GXCECF08"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLP"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI."
BASE COUNT      202 a 127 c 167 g 229 t
ORIGIN
Query Match      83.8%; Score 390.6; DB 9; Length 795;
Best Local Similarity 97.1%; Pred. No. 2.5e-103;
Matches 440; Conservative 0; Mismatches 9; Indels 4; Gaps 4;
Oy 1 atcaagaacatagagttcgggaacatactctcattcattccatcccccaccccaacttta 60
Db 95 ATCAAGAACAATAGAGTTTGGGCAATATATCTTCACTACCTACCCATCCCAATCTTAA 153
Oy 61 ctctactcattcattctcatttgggaacatcagaagatgtgttcgttgaagta 120
Db 154 CTCTACTCATCTCATTCATTAATTTTGGAAATCATCAGAGATGTGTTCGTGAGTA 213
Oy 121 agagattaaagaataagctttttgacccctgcacacacccatgcacaggtgtgtgtac 180
Db 214 AGAGATTAAAGAATAAGCTTTTGGACCCCTGCCAACACCCCACTCCCACTGGTTCAC 273
Oy 181 cctccaatacaataacatgcagaagaagtaagtgtgcccctttctgacggtaactgtgc 241
Db 274 CTTCCAAATCAATAAGATGCCAGGAAGATGAATGTGCTTCTGATGCTGCAATCTCC 313
Oy 241 atcattctccattctccagttctcttccatttcacagtcacactcaggtcagagat 300
Db 334 ATCATCTTCCCATCTTCCAGTCT-CTTTCATTTGCAAGTACAACTCTGGGTCTCAGGAT 392
Oy 301 tataccctctatgtctcgatcattgtcttccactgtgacctgagctgcacctgcac 360
Db 393 TATACCCGCTTATGTCTGGAACATGTCTTTCATCTTGTGCCATCAGCTGGACCTTCCAC 452
Oy 361 ctggagagagtgctcttttgcctcattccctgactccacagaagaacaaagggcagctct 420
Db 453 CT-GGAGAGGGGGCTCTTTGGCTCATCCCTGACTCTCACAAGAACAAGGGCAGCTTCT 511
Oy 421 cactactggacttcaccactttcagggttaag 453
Db 512 CATCATC-GGGCTTCACCACTTTTCAGGGTAAGG 543

```

---

```

RESULT 12
BF511150
LOCUS      523 bp      mRNA      linear      EST 06-DEC-2000
DEFINITION      UI-H-B14-afn-h-06-0-UI-s1 NCI_CGAP_Sub8 Homo sapiens cDNA clone
IMAGE:3087995 3', mRNA sequence.
ACCESSION      BF511150
VERSION        BF511150.1 GI:11594448
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
1 (bases 1 to 523)
NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Unpublished (1997)
Tumor Gene Index
Contact: Robert Strausberg, Ph.D.
Email: cgarbs@mail.nih.gov
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
NCI_CGAP clone distribution information can be found through the
I.M.A.G.E. Consortium/ILNL at:
www.bic.llnl.gov/bbrp/image/image.html The following repetitive
elements were found in this cDNA sequence: 1-22.
>AL_rich:icx.complexity
Seq primer: M13 Forward
POLY-A=yes.
Location/Qualifiers
1..523
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3087995"
/clone_lib="NCI_CGAP_Sub8"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73B-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; NCI_CGAP_Sub8
is a subtracted library derived from NCI_CGAP_Sub5. The
NCI_CGAP_Sub8 library had 2.5 million recombinants. A
single-stranded DNA preparation of NCI_CGAP_Sub5 was used
as a tracer in a subtractive hybridization with a driver
comprising a pool of clones from NCI_CGAP_Sub5 (IMAGE
clone ids 2732833-2737415, 3068040-3069191; 25% of the
driver population), a pool of clones from NCI_CGAP_Sub4
(IMAGE clone ids 2723592-2729326; 25% of the driver
population), NCI_CGAP_Sub6 (pool AIF-AJU, IMAGE ids
2728969-2733190; 25% of the driver population), and
NCI_CGAP_Sub7 (IMAGE ids 3069192-3072238, 3081864-3084550
; 25% of the driver population). Subtraction was
performed as previously described [Bonaldo, Lennon &
Soares (1995): Normalization and Subtraction: Two
Approaches To Facilitate Gene Discovery. Genome Research
6, 751-806.
TAG_LIB=NCI_CGAP_GC4
TAG_TISSUE=derm Cell
TAG_SEQ=AAATC"
BASE COUNT      140 a 130 c 85 g 168 t
ORIGIN
Query Match      83.5%; Score 389.2; DB 10; Length 523;
Best Local Similarity 98.6%; Pred. No. 5.9e-103;
Matches 424; Conservative 0; Mismatches 3; Indels 3; Gaps 3;
Oy 1 utcaagaacatagagttcgggaacatactctcattcattccatcccccaccccaacttta 60
Db 15 ATCAAGAACAATAGAGTTTGGGCAATATATCTTCACTACCTACCCATCCCAATCTTAA 153
Oy 61 ctctactcattcattctcatttgggaacatcagaagatgtgttcgttgaagta 120

```



/tissue\_type="well-differentiated endometrial;  
adenocarcinoma, 7 pooled tumors"  
/lab\_host="DH10B"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.75 kb. Life Technologies catalog #:  
11538-014"

BASE COUNT 166 a 160 c 119 g 200 t  
ORIGIN

Query Match 82.8%; Score 385.8; DB 9; Length 545;  
Best Local Similarity 96.5%; Pred. No. 6.2e-102;  
Matches 437; Conservative 0; Mismatches 12; Indels 4; Gaps 4:

Oy 1 atcaagaacataagattcgggcaataatactctacccatccaccacccaaatttta 60  
|||||  
Db 96 ATCAAGAACAATAGAGTT-GGGCAATATCTTCACTACCCATCCCAACCAATCTTA 154  
Oy 61 ctctact 120  
|||||  
Db 155 CTCTACT 214  
Oy 121 agagattaaaaaataagctttttgacccctgccaacaccccatgcccagagctttac 180  
|||||  
Db 215 AGAGATTAAAGAAATAGCTTTTGGGCAATATCTTCACTACCCATCCCAACCAATCTTA 274  
Oy 181 ctctcaatacaatacatcgaagaagtaagttgacctctctctctctctctctct 240  
|||||  
Db 275 CTCTCAATACAAATAGATGCCAGGAAGTAAAGTTGGCCCTTCTCTCTCTCTCTCT 334  
Oy 241 atcatcttcccatcttccagctctctctctctctctctctctctctctctctct 300  
|||||  
Db 335 ATCATCTTCCCATCTCCAGTCT-CTTTGATTGCAAGTCACAATCTGGGTCTCAAGAT 393  
Oy 301 tataccgctctagttctgaatcatgcttctctctctctctctctctctctctct 360  
|||||  
Db 394 TATACCGCTTAGTCTCGATCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 453  
Oy 361 ctggagagaggtgacctcttctctctctctctctctctctctctctctctctct 420  
|||||  
Db 454 CT-GGAGGAGGTGCGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 512  
Oy 421 catcactggggtctccaccacttccagggttaag 453  
|||||  
Db 513 CATCACT-GGGCTTCACCACTTTCAGGTAAGG 544

RESULT 15  
AI985834  
LOCUS  
DEFINITION  
w78c12.x1 NCI-CGAP\_Ut1 Homo sapiens cDNA clone IMAGE:2493814 3'  
similar to TR:015205 O15205 DIUBIQUitin.. mRNA sequence.  
ACCESSION  
AI985834  
VERSION  
AI985834.1 GI:5813111  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 501)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
Tumor Gene Index  
Unpublished (1997)  
JOURNAL  
Contact: Robert Strausberg, Ph.D.  
Email: [cgaps@mail.nih.gov](mailto:cgaps@mail.nih.gov)  
Tissue Procurement: Christopher Miskaluk, M.D., Ph.D. Michael E.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:  
xxx-bio.llnl.gov/abrp/image/image.html  
Insert length: 678 Std Error: 0.00  
Seq primer: -400P from Gibco  
High quality sequence stop: 421.  
Location/Qualifiers

FEATURES  
Source

1..501  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NCI-CGAP\_Ut1"  
/tissue\_type="well-differentiated endometrial  
adenocarcinoma, 7 pooled tumors"  
/lab\_host="DH10B"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.75 kb. Life Technologies catalog #:  
11538-014"  
BASE COUNT 150 a 154 c 107 g 179 t 1 others  
ORIGIN

Query Match 82.6%; Score 385; DB 9; Length 501;  
Best Local Similarity 97.6%; Pred. No. 1e-101;  
Matches 443; Conservative 0; Mismatches 6; Indels 5; Gaps 5:  
Oy 1 atcaagaacataagattcgggcaataatactctacccatccaccacccaaatttta 60  
|||||  
Db 78 ATCAAGAACAATAGAGTT-GGGCAATATCTTCACTACCCATCCCAACCAATCTTA 136  
Oy 51 ctctact 120  
|||||  
Db 137 CTCTACT 196  
Oy 121 gaaataaagaataagctttttgacccct-gccaacacccatgccaggtggtca 179  
|||||  
Db 137 AGAGATTAAAGAAATAGCTTTTGGGCAATATCTTCACTACCCATCCCAACCAATCTTA 256  
Oy 150 cctctcaatacaatacatgcccagagagtaagttgaccttctctctctctctctct 239  
|||||  
Db 257 CCTCTCAAATACAAATAGATGCCAGGAAGTAAAGTTGCCCTTCTCTGATGCCGTAACTGTC 315  
Oy 241 tataccgctctagttctgaatcatgcttctctctctctctctctctctctctct 299  
|||||  
Db 317 TATACCGCTTAGTCTCGATCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 375  
Oy 350 ttataccgctctagttctgaatcatgcttctctctctctctctctctctctctct 359  
|||||  
Db 376 TATACCGCTTAGTCTCGATCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 435  
Oy 350 cctggagagaggtgaccttctctctctctctctctctctctctctctctctctct 419  
|||||  
Db 476 CCT-GGAGGAGGTGCGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 494  
Oy 420 tcatacactggggtctccaccacttccagggttaag 453  
|||||  
Db 495 TCATCACT-GGGCTTCACCACTTTCAGGTAAGG 527

Search completed: July 29, 2002, 17:20:22  
Job time: 7743 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 17:55:00 ; Search time 1928.75 seconds  
(without alignments)  
3992.703 Million cell updates/sec

Title: US-09-880-107-3847

Perfect score: 368  
Sequence: 1 gaagagacgtgtaagtgcg.....aataatgtatctctatc 369

Scoring table:

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_em.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vl.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_cm.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vl.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_inv.\*

# SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match Length	ID	Description
-----				

1	368	100.0	368	6	AX332609	Sequence
2	368	100.0	368	6	AX333248	Sequence
3	368	100.0	368	9	HSPST1	Y00705 Homo sapien
4	344.8	93.7	608	6	AX014216	Sequence
5	340.8	92.6	432	9	HUMPS11	M11949 Human pancr
6	307.8	83.6	341	6	E01574	cDNA sequen
7	211.4	57.4	265	5	AR166862	Sequence
8	211.4	57.4	265	6	AX192488	Sequence
9	168.8	45.9	179	6	E01725	Synthetic D
10	168.8	45.9	179	6	E02456	DNA encodin
11	152.2	44.2	341	12	SYNECOPST	M22196 Synthetic E
12	152.2	44.1	420	6	AR166863	Sequence
13	152.2	44.1	420	6	AR192489	Sequence
14	162	44.0	302	6	AR059723	Sequence
15	162	44.0	302	6	E08411	DNA sequenc
16	160.8	43.7	234	6	E01819	Sequence
17	160.8	43.7	234	6	E03120	DNA sequenc
18	160	43.5	204	12	SYNPST1B	M11103 Synthetic h
19	159.6	43.4	379	10	RATPST1B	M27883 Rat pancrea
20	158	42.9	377	10	RNMNPEP	X59695 R.norvegicu
21	156.4	42.5	359	10	RATPST1A	M27882 Rat pancrea
22	145.8	39.9	369	10	MMP1MP12	X06342 Mouse mRNA
23	144.6	39.3	238	10	RATPST1A	M35299 Rat pancrea
24	139.8	38.0	317	10	RATMONPA	M22162 Rat monitor
25	119	32.3	422	9	HUMPS1A1	M22971 Human pancr
26	119	32.3	7370	9	AF2866028	AF286028 Homo sapi
27	119	32.3	150255	2	AC011352	AC011352 Homo sapi
28	119	32.3	165889	2	AC011402	AC011402 Homo sapi
29	105.8	29.0	127	9	HUMPS1A3	M20529 Human pancr
30	103.8	28.2	182	6	A11350	Artificial
31	103.8	28.2	182	6	A11352	Artificial
32	103.8	28.2	182	6	A24905	A24905 PSTI master
33	103.8	28.2	182	12	SYNPST1A	M1832 Synthetic h
34	103.8	28.2	185	6	A24907	PSTI master
35	103.8	28.2	5635	12	ASVPSKAN8	X84307 Artificial
36	102.6	27.9	482	6	E03479	E03479 cDNA encod
37	96.4	26.2	289	6	A14318	A14318 c1131-PSTI-
38	96.4	26.2	289	6	A14319	A14319 c1131-PSTI-
39	96	26.1	168	6	I08731	I08731 Sequence 19
40	95	25.8	226	9	HUMPS1A4	M20530 Human pancr
41	92.2	24.2	150123	2	AC011344	AC011344 Homo sapi
42	80.6	21.9	7008	6	AX251285	AX251285 Sequence
43	73.2	19.9	7008	6	AX251286	AX251286 Sequence
44	73.8	17.3	147	10	RATPST13	D11324 Rattus norv
45	63.6	17.3	54541	2	AC104867	AC104867 Mus muscu

## ALIGNMENTS

RESULT 1  
AX332609  
LOCUS AX332609 368 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 3118 from Patent WO0194629.  
ACCESSION AX332609  
VERSION AX332609.1 GI:18123243  
KEYWORDS Human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS 1 (sites)  
TITLE Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,  
Horrigan,S., Soppet,D.R. and Weaver,Z.  
Cancer gene determination and therapeutic screening using signature  
gene sets  
JOURNAL Patent: WO 0194629-A 3118 13-DEC-2001;  
Avalon Pharmaceuticals (US)  
FEATURES  
score Location/Qualifiers  
1..368  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE C N1 100 a 79 c 89 g 100 t

ORIGIN

Query Match 100.0%; Score 368; DB 6; Length 368;  
Best Local Similarity 100.0%; Pred. No. 1.4e-108;  
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gaagagacgtgtaagtgcggtgcggttttcaactgaacctctggagcagagaacttcagcc 60  
|||||  
DB 1 GAAGAGACGTGTAGTGCCTGCGAGTTCACACGACCTCTGGAGCGAGACCTTCAGCC 60  
OY 61 atgaagtaaacaggcatctttcttcagtcgcttggccctcttgaagtcatactgataac 120  
|||||  
DB 61 ATGAGGTAAACAGGCACTTCTCTCAGTCGCTTGGCCCTGTGAGCTATCTCGGTAAAC 120  
OY 121 actgagctgactccctcgggaagagagcccaaatcttacaatgaacttaataatgcagcc 180  
|||||  
DB 121 ACTGGAGCTGACTCCCTCGGAAGAGAGGCGCAAAATGTATCAATGAACCTTAATGSAATGCACC 180  
OY 181 aagatatagacctgtctgtgggactatgaaataacttcccaatgaatgcctgatta 240  
|||||  
DB 181 AGATATATGACCCCTGCTGCGGACTGATGGAATACTATCCCAATGAATGCGGTGTTA 240  
OY 241 ttttttgaaggtcggaacgcgcagactctctctcattcaaaaactcgggcttgcctga 300  
|||||  
DB 241 TGTITTTGAAGGTTCGGAACGCCAGACTTCTATCTCTCATTCATCAAAAATCTGGGCTTGCCTGA 300  
OY 301 gaaccaaggttttgaatacccatcaggtcacccgagagccctatttgaataaaatgtatc 360  
|||||  
DB 301 GAACCAAGGTTTGAATACTCCATCAGGTACCGCGAGGCCATTGTTGAATAAATGATATC 360  
OY 361 tgaatc 368  
|||||  
DB 361 TGAATAIC 368

RESULT 2

AX333248  
LOCUS AX333248 368 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 3757 from Patent WO0194629.  
ACCESSION AX333248  
VERSION AX333248.1 GI:18123882  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (sites)  
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,  
Horrikan,S., Soppet,D.R. and Weaver,Z.  
TITLE Cancer gene determination and therapeutic screening using signature  
gene sets  
JOURNAL Patent: WO 0194629-A 3757 13-DEC-2001;  
Avalon Pharmaceuticals (US)  
FEATURES  
source  
1..368  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 100 a 75 c 89 g 100 t  
ORIGIN

Query Match 100.0%; Score 368; DB 6; Length 368;  
Best Local Similarity 100.0%; Pred. No. 1.4e-108;  
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gaagagacgtgtaagtgcggtgcggttttcaactgaacctctggagcagagaacttcagcc 60  
|||||  
DB 1 GAAGAGACGTGTAGTGCCTGCGAGTTCACACGACCTCTGGAGCGAGACCTTCAGCC 60  
OY 61 atgaagtaaacaggcatctttcttcagtcgcttggccctcttgaagtcatactgataac 120  
|||||  
DB 61 ATGAGGTAAACAGGCACTTCTCTCAGTCGCTTGGCCCTGTGAGCTATCTCGGTAAAC 120

OY 121 actgagctgactccctcgggaagagagcccaaatgttacaatgaacttaatgagtcagcc 180  
|||||  
DB 121 ACTGAGCTGACTCCCTCGGAAGAGAGGCGCAAAATGTATCAATGAACCTTAATGGAATGCACC 180  
OY 181 aagatatagacctgtctgtgggactatgaaataacttcccaatgaatgcctgatta 240  
|||||  
DB 181 AGATATATGACCCCTGCTGCGGACTGATGGAATACTATCCCAATGAATGCGGTGTTA 240  
OY 241 ttttttgaaggtcggaacgcgcagactctctcattcaaaaactcgggcttgcctga 300  
|||||  
DB 241 TGTITTTGAAGGTTCGGAACGCCAGACTTCTATCTCTCATTCATCAAAAATCTGGGCTTGCCTGA 300  
OY 301 gaaccaaggttttgaatacccatcaggtcacccgagagccctatttgaataaaatgtatc 360  
|||||  
DB 301 GAACCAAGGTTTGAATACTCCATCAGGTACCGCGAGGCCATTGTTGAATAAATGATATC 360  
OY 361 tgaatc 368  
|||||  
DB 361 TGAATAIC 368

RESULT 3

AX333248  
LOCUS AX333248 368 bp mRNA linear PRI 12-SEP-1993  
DEFINITION Homo sapiens pst1 mRNA for pancreatic secretory inhibitor  
(expressed in neoplastic tissue).  
ACCESSION Y00705  
VERSION Y00705.1 GI:35765  
KEYWORDS pst1 gene; trypsin inhibitor.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 368)  
AUTHORS Tomita,N.  
TITLE Direct Submission  
JOURNAL Submitted (15-NOV-1987) Tomita N., Institute of Molecular and  
Cellular Biology, Osaka University, Yamada-oka, Suita 565, Japan  
REFERENCE 2 (bases 1 to 368)  
AUTHORS Tomita,N., Horii,A., Yamamoto,T. and Oawara,M.  
TITLE Expression of pancreatic secretory trypsin inhibitor (PSTI) gene in  
neoplastic tissues  
JOURNAL F595 Lett. (1987) In press  
COMMENT This sequence is identical with that of PSTI RNA <M11949> isolated  
from human pancreatic cDNA library.  
FEATURES  
Location/Qualifiers  
1..368  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="lambda TI-Cl"  
/tissue\_type="sigmoid colon"  
51..300  
/note="inhibitor (AA 1-79)"  
/codon\_start=1  
/protein\_id="CAA68697.1"  
/db\_xref="GI:35765"  
/db\_xref="SWISS-PROT:P00995"  
/translation="MKVIGIFLLSALLSLSGNTGADSLGKRAKYNELNGCTKIYD  
FVCGIDGNTYPRVCVLCFEGKRRKQTSILIKSGPC"  
349..354  
/note="polyA signal"  
polyA\_site 358  
/note="polyA site"  
BASE COUNT 100 a 79 c 89 g 100 t  
ORIGIN

CS

Query Match 100.0%; Score 368; DB 9; Length 368;  
Best Local Similarity 100.0%; Pred. No. 1.4e-108;  
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 aaagagacgtgtaagtgcggtgcggttttcaactgaacctctggagcagagaacttcagcc 60  
|||||  
DB 1 GAAGAGACGTGTAGTGCCTGCGAGTTCACACGACCTCTGGAGCGAGACCTTCAGCC 60  
OY 61 atgaagtaaacaggcatctttcttcagtcgcttggccctcttgaagtcatactgataac 120  
|||||  
DB 61 ATGAGGTAAACAGGCACTTCTCTCAGTCGCTTGGCCCTGTGAGCTATCTCGGTAAAC 120

```

Db 1 GAAGAGAGGTCGTAAGTCGGTGGAGTTCGAACTGACCTGGAGCGAGAACATTCAGCC 60
Oy 61 atgaaggaagcagcagcattcttctcaagtcgcttgagccctcttgaactctatctogtaac 120
Db 61 ATGAGGTAACAGGATCTTCTTCAGTCAGTCGCTTGGCCCTTGAGTCTATCTGGTAAC 120
Oy 121 actgaagtcgactcctcctgggaagcagagagcccaaatgttacaatgaacttaattggtatgcacc 180
Db 121 ACTGGAGTGTACTCTCGTGGAGAGAGGCGCAATGTTACAAATGAACATAATGCGATGCACC 180
Oy 181 aagatatatgacctgtctgtggaactatgataaataactctatcccaatgaatgcggtgtta 240
Db 181 AGATAATATGACCTCTGTGGGACTGATGGAATATCTATCCCAATGAATGAAATCGTGTTA 240
Oy 241 tgtttgaagtcggaagcgcagactctctatcctcattcacaatactcgggcccctgtgtga 300
Db 241 TGTTTGAAGTTCGGAACGCCAGACTCTATCTCTATTCATTCATAAATCTGGTTCGTA 300
Oy 301 gaaccaagttttgaaatcccatcaggtcaccgcagagccctattgttgaatgaatgatatc 360
Db 301 GAACCAAGGTTTGAATCCCATCAGGTCAACCGGAGGCGCTATTGTTGAAATAATGATATC 360
Oy 361 tgaatatac 368
Db 361 TGAATATC 368

RESULT 4
LOCUS AX014216 608 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 11 from Patent WO954445.
ACCESSION AX014216
VERSION AX014216.1 GI:10040617
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 608)
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
Pilarczyk,C.
TITLE Human nucleic acid sequences from normal pancreas tissue
JOURNAL Patent: WO 954445-A 11 28-OCT-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); MEIAGEN GES FUER GENOMFORSCHUN
(DE); PILARSKY CHRISTIAN (DE)
FEATURES
Source
Location/Qualifiers
1..608
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 182 a 128 c 148 g 150 t
ORIGIN
Query Match 93.7% Score 344.8; DB 6; Length 608;
Best Local Similarity 96.8% Pred. No. 5,le-101;
Matches 365; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

Oy 1 caagagacgtatgtagcagtcagtttcaactgaacctgaacctcagacacagaacttcagcc 60
Db 204 CAAGAGACGTGTAAGTCGGTGGAGTTCGAACTGACCTGGAGCGAGAACATTCAGCC 263
Oy 61 atgaaggaagcagcagcattcttctcaagtcgcttgagccctcttgaactctatctogtaac 120
Db 264 ATGAGGTAACAGGATCTTCTTCAGTCAGTCGCTTGGCCCTTGAGTCTATCTGGTAAC 323
Oy 121 actgaagtcgactcctcctgggaagcagagagcccaaatgttacaatgaacttaattggtatgcacc 180
Db 121 ACTGGAGTGTACTCTCGTGGAGAGAGGCGCAATGTTACAAATGAACATAATGCGATGCACC 383
Oy 181 aagatatatgacctgtctgtggaactatgataaataactctatcccaatgaatgcggtgtta 240

```

```

Db 394 AAGATAATGACCCCTGTCGTGGAGTGTGGAATACTTATCCCAATGATCGCTGTTA 443
Oy 241 tgttttzaaggtcagaaacccagacttctatcctcattcaaaaactcggccctgtctga 300
Db 444 TGTTCGAAATGCGGAACGCCAGACTTCTATCTCTATTCATAAATCTGGGCCCTTGCTGA 503
Oy 301 gaaccaaggttttgaatacccatcagatcaccgcagagcc-----tattgttga 350
Db 504 GAACCAAGGTTTGAATCCCATCAGGTCAACCGGAGGCGCTGACTGSCCTTATGTIGAA 563
Oy 351 taaatgtatctatgaalatac 368
Db 564 TAAATGTAICTGAATATC 581

RESULT 5
LOCUS HUMPSII 432 bp mRNA linear PRI 27-APR-1993
DEFINITION Human pancreatic secretory trypsin inhibitor (PSII) mRNA, complete cds.
ACCESSION M11949
VERSION M11949.1 GI:190687
KEYWORDS
SOURCE Human pancreas cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 432)
AUTHORS Yamamoto,T., Nakamura,Y., Nishide,T., Emi,M., Ogawa,M., Mori,T. and
Matsubara,K.
TITLE Molecular cloning and nucleotide sequence of human pancreatic
secretory trypsin inhibitor (PSII) cDNA
JOURNAL Biochem. Biophys. Res. Commun. 132, 605-612 (1995)
MEDLINE 96050645
FEATURES
Location/Qualifiers
1..432
/organism="Homo sapiens"
/db_xref="taxon:9606"
117..355
/notes="pancreatic secretory trypsin inhibitor"
/codon_start=1
/protein_id="AAA36521.1"
/db_xref="GI:190688"
/translation="MKVTGIFLALALLSLGNTGADSLGPEAKCYNELNGCTIYD
PCGTDGNTYPNECVLCFENPKRQTSILIQKSPC"
BASE COUNT 121 a 96 c 103 g 112 t
ORIGIN
Query Match 92.6% Score 340.8; DB 9; Length 432;
Best Local Similarity 96.8% Pred. No. 1e-99;
Matches 362; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

Oy 3 agagacgtatgaagtcagtcagtttcaactgaacctcagacacagacacttcagccat 62
Db 59 AGAGACGTGTAAGTCGGTGGAGTTCGAACTGACCTGGAGCGAGAACATTCAGCCAT 118
Oy 63 aaagtaaacagcactcttctcagtcgcttgagccctctgaagtcctatctcagtaaac 122
Db 119 GAAGGTAACAGGATCTTCTTCAGTCGCTTGGCCCTTGAGTCTATCTGGTAACAC 178
Oy 123 tgaagtcgactcctcctgggaagcagagcccaaatgttacaatgaacttaattggtatgcacc 182
Db 123 TGAAGTGTACTCTCGTGGAGAGAGGCGCAATGTTACAAATGAACATAATGCGATGCACC 238
Oy 183 aatatacactcctctctggaactatgataaataactctatcccaatgaatgcggtgtta 242
Db 239 ATATATGACCTCTGTGGGACTGATGGAATATCTATCCCAATGAATGCGTGTGTTATG 299
Oy 243 ttttgaagtcggaagcagcagactctctatcctcattcacaatactcggcccttgctgaa 302
Db 243 TTTGAAATTCGGAACGCCAGACTTCTATCTCTATTCATAAATCTGGGCCCTTGCTGACA 358

```

```

QY 303 accaaggttttgaatcccatcagtcacccgcgagcc-----tattgttgaata 352
|||||
DB 359 ACCAAGGTTTGAATCCCATCAGGTACCGCGAGGCTGACTGGCCITATGTTGAATA 418
|||||

QY 353 aatgtatctgaata 366
|||||
DB 419 AATGTAATCTGAATA 432
|||||

RESULT 6
E01574
LOCUS CDNA sequence of human PST1. 341 bp RNA linear PAT 29-SEP-1997
ACCESSION E01574
VERSION E01574.1 GI:2169827
KEYWORDS JP 1988098397-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 341)
AUTHORS Ogawa,M. and Matsubara,K.
TITLE PRODUCTION OF HUMAN PST1 BY YEAST
JOURNAL Patent: JP 1988098397-A 1 28-APR-1989;
SHIONOGI & CO LTD
COMMENT OS human
PN JP 1988098397-A/1
PD 28-APR-1988
PF 14-OCT-1986 JP 1986245049
PI OGAWA MICHIO, MATSUBARA KENICHI
PC C12P21/02.C12N15/00//C12N9/99.(C12P21/02.C12R1:645): CC
strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: tissue_type=pancreas;
FH Key Location/Qualifiers
FT 5'UTR 1..25
FT CDS 26..265
FT 3'UTR
FT 266..341
FT sig_peptide 26..94
FT mat_peptide 95..262
FT /product='human PST1 precursor' FT 3'UTR
FT /product='human PST1'
FEATURES
source
1..341
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 94 a 75 c 78 g 93 t
ORIGIN

Query Match 83.6%; Score 307.8; DB 6; Length 341;
Best Local Similarity 95.5%; Pred. No. 5.3e-89;
Matches 329; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

QY 36 gactcttgagcaggaacttcagccatcagggtaacagggatctttcttcagtgccct 95
|||||
DB 1 GACCTCTGGAGCGAGAACITTCAGGCAITGAAGGTAAACAGGCAITCTTCTCAGTGCCT 60
|||||

QY 96 gcccctgttgagtcctatctgttaactgaagctgacctccctgggaagagagggccaatg 153
|||||
DB 61 GCCCTGTGTGAGTCTAICTGTGTAACACTGGAGCTGACTCCCTGGGAAGAGAGGCCAATG 120
|||||

QY 156 ttacaatgaactaatggaatcaccagaatataatgacctgtctgttgaggactgagaa 215
|||||
DB 121 TTACAATGAACCTTAATGGATGCACCAAGATATATGACCCCTGTCTGTGGGACGTAIGGAA 180
|||||

QY 216 tacttatcccaataacatcgcttatctattttgaagtcgaaacgccagacattctatct 273
|||||
DB 181 TACTTATCCCAATGAATCGCTGTGTTAIGTTTTTGAATAATCGGNAACGCCAGACTTCACT 240
|||||

```

```

QY 216 catlcaaaaatctgggcttctcagaaacaaagttttgaaatcccatcaggtaaccgcg 335
|||||
DB 241 CATTCAAAATCTGGGCTTCTCTGAGAACCAAGGTTTTTGAATCCCACTACGTCACCGCG 300
|||||

QY 336 aggc-----tattgttgaataaatgtatctgaata 366
|||||
DB 301 AGGCCTAGACGGCTTATGTTGAATAAATGATCTGAATA 341
|||||

RESULT 7
E0156862
LOCUS ARI56862 265 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 55 from patent US 6284241.
ACCESSION ARI56862
VERSION ARI56862.1 GI:16243230
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 265)
AUTHORS Xu,J.
TITLE Compounds for immunotherapy and diagnosis of colon cancer and
methods for their use
JOURNAL Patent: US 6284241-A 55 04-SEP-2001;
FEATURES Location/Qualifiers
1..265
source
/organism='unknown'
BASE COUNT 69 a 59 c 57 g 75 t 5 others
ORIGIN

Query Match 57.4%; Score 211.4; DB 6; Length 265;
Best Local Similarity 95.6%; Pred. No. 1.1e-57;
Matches 237; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 76 atctttcttcagtcacctgacctgttgagctatctgttaacacacgtggagctaacctc 135
|||||
DB 1 ATCTTCTTCAGTGCCTTGGCCNTTGTGAGTCTATCTGTTAACACACGTGGAGCTGACCTC 60
|||||

QY 136 ctgggaagagagggccaaatgttacaatgaacttaataatgaatgacccaataatatagacct 195
|||||
DB 61 CTGGGAAGAGAGGCCAAATGTTACAATGAACCTTAATGGATGCACCAAGATATATAGCCCT 120
|||||

QY 196 atctatgagactgaggaataactatcccaatgaatg-cgtgttatg-ttttgaagtc 253
|||||
DB 121 GTCTGGGAGCTGATGGAAATACCTTATCCCAATGAATGCCGTGTTATGTTTGAATAATC 180
|||||

QY 253 ggaacgcccagactctctatccctcatcgaataatcgggacctgtggaacacaaagtttt 313
|||||
DB 181 GGAACGCCAGACTTCTATCCCTCAITCAAAAATCTGGGCCCTTCTGAAACCCAGGGTTT 240
|||||

QY 314 gaaatccc 321
|||||
DB 241 NAAATCC 248
|||||

RESULT 8
E0192488
LOCUS AX192488 265 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequences 55 from Patent WO0149716.
ACCESSION AX192488
VERSION AX192488.1 GI:15210452
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 265)
AUTHORS Xu,J., Lodes,M.J., Secret,H., Benson,D.R., Meagher,M.J.,
Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.
TITLE Compounds for immunotherapy and diagnosis of colon cancer and
methods for their use
JOURNAL Patent: WO 0149716-A 55 12-JUL-2001;

```

CORIXA CORPORATION (US)		Location/Qualifiers			
source		1..265			
BASE COUNT		69 a 59 c 57 g 75 t 5 others			
ORIGIN					
Query Match		57.4%; Score 211.4; DB 5; Length 245;			
Best Local Similarity		95.6%; Pred. No. 1.1e-57;			
Matches 237; Conservative		0; Mismatches 9; Indels 2; Gaps 2;			
Qy	76	atctttctcgaatgccctggccctgttggaatctatctggtataacacactggagctgactcc	135		
Db	1	ATCTTCTTCAGTGGCCCTGGCCCTGTGAGTCTATCTGTTGAACACTGGAGCTGACTCC	60		
Qy	136	ctgggaagagagcccaaatgttacaatgaacttaagatgcaccaagatatatgacct	195		
Db	61	CTGGGAAGAGAGGCCCAATGTTACAATGAACITTAATGATGCACCAAGATAATGACCT	120		
Qy	196	gtctgtgggactgaatgaataacttatcccaatgaatgcgtgttatg-ttttgaagctc	253		
Db	121	GTCGTGGGACTGATGGAATACTTATCCCAATGAATGCCGTGTATGTTTITGAAATC	180		
Qy	254	ggaaacccagactctatccctcattcaaaaatctggacccttgctgaagaccaggtttt	313		
Db	181	GGAAACCCAGACTCTATCCCTATTCATCAAAAATCTGGCCCTTNCIGRAAACAGGGTIT	240		
Qy	314	gaatccc 321			
Db	241	NAAAATCC 248			
RESULT 9					
E01725	E01725		179 bp	DNA	linear
LOCUS	Synthetic DNA encoding human pancreatic secretory trypsin inhibitor (PSII).				PAT 29-SEP-1997
DEFINITION	E01725				
ACCESSION	E01725				
VERSION	E01725.1 GI:2169978				
KEYWORDS	JP 1988267289-A/1.				
SOURCE	synthetic construct.				
ORGANISM	artificial sequence.				
REFERENCE	1 (bases 1 to 179)				
AUTHORS	Obara.O., Shin.M., Kikuchi.N. and Teraoka.H.				
TITLE	NOVEL PRODUCTION OF PROTEIN				
JOURNAL	Patent: JP 1988267289-A I 04-NOV-1988;				
COMMENT	SHIONOGI & CO LTD				
OS	Artificial gene				
OC	Artificial sequence: Genes.				
PN	JP 1988267289-A/1				
PD	04-NOV-1988				
PF	05-OCT-1987 JP 1987253437				
PR	14-OCT-1986 JP 86P 245048, 23-DEC-1986 JP 86P 314503 PI				
OBARA OZAMU, SHIN MASARU, KIKUCHI NORIHISA, TERAOKA HIROSHI FC	C12P21/00.C12N15/00.(C12P21/00.C12R1:19):				
CC	strandedness: Double;				
CC	topology: Linear;				
CC	hypothetical: No;				
CC	anti-sense: No;				
EH	Key				
EH	Location/Qualifiers				
FI	CDS				
FI	mat_peptide				
FI	Location/Qualifiers				
1..179	/product='synthetic human psi-1'				
source					
BASE COUNT	54 a 37 c 41 g 47 t				

CORIXA CORPORATION (US)		Location/Qualifiers			
source		1..265			
BASE COUNT		69 a 59 c 57 g 75 t 5 others			
ORIGIN					
Query Match		57.4%; Score 211.4; DB 5; Length 245;			
Best Local Similarity		95.6%; Pred. No. 1.1e-57;			
Matches 237; Conservative		0; Mismatches 9; Indels 2; Gaps 2;			
Qy	76	atctttctcgaatgccctggccctgttggaatctatctggtataacacactggagctgactcc	135		
Db	1	ATCTTCTTCAGTGGCCCTGGCCCTGTGAGTCTATCTGTTGAACACTGGAGCTGACTCC	60		
Qy	136	ctgggaagagagcccaaatgttacaatgaacttaagatgcaccaagatatatgacct	195		
Db	61	CTGGGAAGAGAGGCCCAATGTTACAATGAACITTAATGATGCACCAAGATAATGACCT	120		
Qy	196	gtctgtgggactgaatgaataacttatcccaatgaatgcgtgttatg-ttttgaagctc	253		
Db	121	GTCGTGGGACTGATGGAATACTTATCCCAATGAATGCCGTGTATGTTTITGAAATC	180		
Qy	254	ggaaacccagactctatccctcattcaaaaatctggacccttgctgaagaccaggtttt	313		
Db	181	GGAAACCCAGACTCTATCCCTATTCATCAAAAATCTGGCCCTTNCIGRAAACAGGGTIT	240		
Qy	314	gaatccc 321			
Db	241	NAAAATCC 248			
RESULT 9					
E01725	E01725		179 bp	DNA	linear
LOCUS	Synthetic DNA encoding human pancreatic secretory trypsin inhibitor (PSII).				PAT 29-SEP-1997
DEFINITION	E01725				
ACCESSION	E01725				
VERSION	E01725.1 GI:2169978				
KEYWORDS	JP 1988267289-A/1.				
SOURCE	synthetic construct.				
ORGANISM	artificial sequence.				
REFERENCE	1 (bases 1 to 179)				
AUTHORS	Obara.O., Shin.M., Kikuchi.N. and Teraoka.H.				
TITLE	NOVEL PRODUCTION OF PROTEIN				
JOURNAL	Patent: JP 1988267289-A I 04-NOV-1988;				
COMMENT	SHIONOGI & CO LTD				
OS	Artificial gene				
OC	Artificial sequence: Genes.				
PN	JP 1988267289-A/1				
PD	04-NOV-1988				
PF	05-OCT-1987 JP 1987253437				
PR	14-OCT-1986 JP 86P 245048, 23-DEC-1986 JP 8				

```

Query Match          45.9%   Score 168.8;   DB 6;   Length 179;
Best Local Similarity 28.8%;   Pred. No. 7.9e-44;
Matches 170; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 130 gactccctgggaagagggccaaaglttacaataaacttaataatgacaccacaaatatat 183
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8 GACTCCCTGGGAAGAGAGGGCCAAATGTTACAACTTAATGAATGACCAAGATATAT 6-

QY 190 gacctctctgtggaactgaataactattcccaataatcgtatattttttaa 242
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68 GACCCTGCTGTGGGACTGTGAGAAATACITATCCCAATGAATCGGTGTATGTTTGA 12-

QY 250 ggtcgaaacccagactctctctcattcaaaaatctggccttgcctgaag 301
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 128 RATCGAAACCCAGACTTCTATCCCTATTCATTCATTCATTCATTCATTCATTC 179

RESULT 11
SYNECOPST
LOCUS          341 bp      DNA      linear      SYN 27-APR-1993
DEFINITION    Synthetic E.coli alkaline phosphatase signal peptide/human
               pancreatic secretory trypsin inhibitor protein (PSII) fusion
               protein gene 5' end.
ACCESSION     M22196.1 GI:208314
VERSION       M25727
KEYWORDS      alkaline phosphatase; fusion protein; pancreatic trypsin inhibitor.
SOURCE        Synthetic DNA.
ORGANISM      synthetic construct
               artificial sequence.
REFERENCE     1 (bases 103 to 341)
               Ogino,H.
AUTHORS       Kanamori,T., Mizushima,S., Shimizu,Y., Morishita,H., Kubota,H.,
               Nii,A., Ogino,H., Nagase,Y., Kisaragi,M. and Nobuhara,M.
JOURNAL       Expression and excretion of human pancreatic secretory trypsin
               inhibitor in lipoprotein-deletion mutant of Escherichia coli
AUTHORS
TITLE         Gene 65, 295-300 (1988)
JOURNAL       89005270
MEDLINE
COMMENT       Draft entry and computer-readable sequence for [2] kindly provided
               by H. Ogino, 21-SEP-1988.
FEATURES
   source      1..341
               Location/Qualifiers
               /organism="synthetic construct"
               /db_xref="taxon:32630"
   sig_peptide 25..88
               /note="alkaline phosphatase signal peptide"
   CDS         25..106
               /note="synthetic fusion protein"
               /codon_start=1
               /protein_table=11
               /protein_id="AA072695.1"
               /db_xref="GI:208315"
               /translation="MKQSTIALALLFLFPTVTKADSLGF"
   mat_peptide 89..103
               /product="pancreatic secretory trypsin inhibitor"
   BASE COUNT 104 a 82 c 57 g 88 t
   ORIGIN      HindIII site.

Query Match          44.2%   Score 162.6;   DB 12;   Length 341;
Best Local Similarity 92.4%;   Pred. No. 8.5e-42;
Matches 171; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 121 actggagctgactccctgggaagagggccaaatttacaataaacttaataatgagacc 182
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 157 ACAAAAGCCGACTCCCTGAGTGGGAGGCGCAATGTTACAACTTAATGATGATGAC 245

QY 181 aagatatagacctgtctgtgggactgatgaaatactattcccaataatgactgttga 242
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 217 AAGATATAGACCCGTCTGTGGGACTGATGGAATATCTATCCCAATGAATGCTGTGTA 2-

```

```

QY 241 tttttgaaggtcggaacgcccagacttctatcctcattcaaaaatctggccttgcgtga 300
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 277 TGTITIGAAATCGGAAGCCAGACATCGATCCCTCATTCATTCATTCATTCATTCAT 336

QY 301 gaacc 305
      |||
Db 237 GGATC 341

RESULT 12
LOCUS          420 bp      DNA      linear      PAT 17-OCT-2001
DEFINITION    AR156863
Sequence 56 from patent US 5284241.
ACCESSION     AR156863
VERSION       AR156863.1 GI:16243231
KEYWORDS      Unknown.
SOURCE        Unknown.
ORGANISM      Unclassified.
               1 (bases 1 to 420)
REFERENCE     Xu,J.
AUTHORS       Compounds for immunotherapy and diagnosis of colon cancer and
               methods for their use
TITLE         Patent: US 5284241-A 56 04-SEP-2001;
               Location/Qualifiers
JOURNAL
FEATURES
   source      1..420
               /organism="unknown"
   BASE COUNT 103 a 107 c 88 g 111 t 11 others
   ORIGIN

Query Match          44.1%   Score 162.2;   DB 6;   Length 420;
Best Local Similarity 96.2%;   Pred. No. 1.2e-41;
Matches 175; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 160 aatgaacttaataatgagatga-ccaagatatatgacctgtctgtgggactgatgaaatac 218
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 210 AANGAACTTAATGATGCCACCCCAAGATATATGATGATGATGATGATGATGATGATG 141

QY 215 ttatcccaataatgactgtattgtttttaaagtcgaaacccagacttctatctcat 278
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111 TAICCCAAATGAATGGGTGTATGTTTTTGAANAATCGGAACGCCAGACTTCTATCTCAT 81

QY 273 tcaaaaatctgggctgtgctgagaaacccaagggttttgaatacccatcagatcccgaggg 338
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 80 TCAAAAATCTGGGCTGTGCTGAGAACCAAGGTTTTTGAATCCCATCAGGTACCCGCGAGG 21

QY 333 cc 340
      |
Db 20 AC 19

RESULT 13
LOCUS          420 bp      DNA      linear      PAT 15-AUG-2001
DEFINITION    AX192482
Sequence 56 from Patent WO0149716.
ACCESSION     AX192489
VERSION       AX192489.1 GI:15210453
KEYWORDS      human.
SOURCE        Homo sapiens.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 420)
AUTHORS       Xu,J., Lodes,M.J., Secret,H., Benson,D.R., Meagher,M.J.,
               Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.
TITLE         Compounds for immunotherapy and diagnosis of colon cancer and
               methods for their use
JOURNAL       Patent: WO 0149716-A 56 12-JUL-2001;
               CORIAX CORPORATION (US)
               Location/Qualifiers
FEATURES
   source      1..420
               /organism="Homo sapiens"

```

BASE COUNT	103 a	107 c	88 g	111 t	11 others
ORIGIN	/db_xref="taxon:9506"				
Query Match	44.0%	Score 162.2	DB 6	Length 420	
Best Local Similarity	95.2%	Pred. No. 1.2e-41			
Matches 175	Conservative	0	Mismatches	6	Indels 1
Gaps	1				
Qy	150	aatgaacttaattgagatgaca-ccaagatatatgacccctgtctgtggaactgaatgaataac	218		
Db	200	AANGAACTTAATGATGACCAACCAAGATATATGACCCNGTCTGTGGGACTGATGGAAATAC	141		
Qy	219	ttatcccaatgaatgcgtgttatgtttttaaagtcgaaacgacacactcttatctctcat	278		
Db	140	TTATCCCAATGAATGCGTGTATGTTTGGAAATCGGAACGCCAGACTTCTATCNICAT	81		
Qy	279	tcaaaaaatctggccctctgtgagacccaaggtttttaaatacccatcaggtcaccccaagg	338		
Db	80	TCAAAAATCTGGGCCCTTCTGTGAGAACCAAGGTTTGGAAATCCCATCAGGTCACCCGAAAG	21		
Qy	339	cc 340			
Db	20	AC 19			
RESULT 14					
AP059723					
LOCUS	AP059723	Sequence 26	from patent US 5840518	linear	FBI 29-SEP-1999
DEFINITION	Sequence 26 from patent US 5840518				
ACCESSION	AP059723				
VERSION	AP059723.1 GI:5986173				
KEYWORDS	Unknown				
SOURCE	Unknown				
ORGANISM	Unclassified				
REFERENCE	1 (bases 1 to 302)				
AUTHORS	Morishita,H., Kanamori,I. and Nobuhara,M.				
TITLE	DNA fragment, vector containing the DNA fragment, transformant transformed with the vector and process for producing protein using the vector				
JOURNAL	Patent: US 5840518-A 26 24-NOV-1998				
FEATURES	source				
BASE COUNT	97 a	64 c	58 g	73 t	
ORIGIN	1..302				
Query Match	44.0%	Score 162	DB 6	Length 302	
Best Local Similarity	91.9%	Pred. No. 1.3e-41			
Matches 171	Conservative	0	Mismatches	15	Indels 0
Gaps	0				
Qy	120	cactggagctgactccctctggaagagagggcccaaatgtttacaaatgaacttaattgagtcac	179		
Db	116	CTCGGAATGGACTCCCTAGGTCCGGAGGCCAAATGTTACAAATGAACCTTAAATGGAATGAC	175		
Qy	180	caagatatgaactctctgtggaactgagaaatacttatcccaatgaatgcgtgtt	239		
Db	176	CAAGATATAIGACCCCTGTCTGTGGGACTGATGGAATACTTATCCCAATGAATGCGTGT	235		
Qy	240	atgttttgaagtcgaaacgacacactctctctcattccaaataatctgtggtg	299		
Db	236	ATGTTTGGAAATCGGAACGCCACATCGATCCTCTATTCATAAATCTGGGCTTGCTG	295		
Qy	300	agaacc 305			
Db	296	AGGATC 301			
RESULT 15					
E08411					
LOCUS	E08411	Sequence 302	bp	DNA	linear FBI 29-SEP-1997

Search completed: July 29, 2002, 17:55:02  
Job time: 5168 sec

DEFINITION	DNA sequence coding signal and target fused protein.				
ACCESSION	E08411				
VERSION	E08411.1 GI:2176528				
KEYWORDS	JP 1994315386-A/10.				
SOURCE	unidentified.				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 302)				
AUTHORS	Morishita,H., Kanamori,I. and Nobuhara,M.				
TITLE	DNA FRAGMENT, VECTOR CONTAINING THE SAME FRAGMENT, TRANSFORMANT TRANSFORMED BY THE SAME VECTOR AND PRODUCTION OF PROTEIN USING THE SAME VECTOR				
JOURNAL	Patent: JP 1994315386-A 10 15-NOV-1994;				
COMMENT	MOCHIDA PHARMACEUT CO LTD				
CS	None				
OC	Artificial sequences.				
FN	JP 1994315386-A/10				
FD	15-NOV-1994				
PF	01-MAY-1993 JP 1993128528				
PI	MORISHITA HIDEAKI, KANAMORI TOSHIYUKI, NOBUHARA MASAHIRO PC				
CI	C12N15/31.C07K13/00.C12N1/21.C12P21/02.C12N15/31.C12R1/19). PC				
FC	C12N1/21). (C12P21/02.C12R1/19):				
CC	strandedness: Single;				
CC	topology: Linear;				
EH	Key				
EH	Location/Qualifiers				
FI	source				
FI	1..302				
FI	/strain="GC4670(pM474)";				
FI	sig_peptide 27..122				
FI	mat_peptide 123..293				
FI	/product="pancreatic secretary trypsin FT inhibitor";				
FEATURES	Location/Qualifiers				
source	1..302				
BASE COUNT	97 a	64 c	58 g	73 t	
ORIGIN	1..302				
Query Match	44.0%	Score 162	DB 6	Length 302	
Best Local Similarity	91.9%	Pred. No. 1.3e-41			
Matches 171	Conservative	0	Mismatches	15	Indels 0
Gaps	0				
Qy	120	cactggagctgactccctctggaagagagggcccaaatgtttacaaatgaacttaattgagtcac	179		
Db	116	CTCGGAATGGACTCCCTAGGTCCGGAGGCCAAATGTTACAAATGAACCTTAAATGGAATGAC	175		
Qy	180	caagatatgaactctctgtggaactgagaaatacttatcccaatgaatgcgtgtt	239		
Db	176	CAAGATATAIGACCCCTGTCTGTGGGACTGATGGAATACTTATCCCAATGAATGCGTGT	235		
Qy	240	atgttttgaagtcgaaacgacacactctctcattccaaataatctgtggtg	299		
Db	236	ATGTTTGGAAATCGGAACGCCACATCGATCCTCTATTCATAAATCTGGGCTTGCTG	295		
Qy	300	agaacc 305			
Db	296	AGGATC 301			

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on:

July 29, 2002, 16:28:54 : Search time 1928.76 seconds  
(without alignments)  
5055.977 Million cell updates/sec

Title: US-09-880-107-2492

Perfect score: 466  
Sequence: 1 atcaagaacacatagagttc.....gggtaaggtggga:ggctctt 466

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3555312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

GenEmbl:

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.or.\*

21: em.ov.\*

22: em.pat.\*

23: em.ph.\*

24: em.pl.\*

25: em.ro.\*

26: em.sts.\*

27: em.un.\*

28: em.vi.\*

29: em.htg.hum.\*

30: em.htg.in.\*

31: em.htg.in.\*

32: em.htg.other.\*

33: em.htg.in.\*

#### Result

No. Score Match Length DB ID

Description

#### SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

1	466	100.0	466	5	AX329950
2	400.2	85.5	155874	2	AL645936
3	398.6	85.5	971	9	AF123050
4	398.6	85.5	134292	9	HS271M21
5	398.6	85.5	211591	2	AL662826
6	397	85.2	777	9	HS21UB1QU
7	397	85.2	41059	2	AC004179
8	397	85.2	129567	2	AL671618
9	397	85.2	129806	9	AC006137
10	395.4	84.8	791	9	BC012472
11	394.4	71.8	474	6	AX237229
12	394.4	71.8	474	6	AX237229
13	323.4	69.4	453	6	AX237670
14	321.4	62.5	6355	10	AF314088
15	151.4	32.5	154614	10	MM573K1
16	142	30.5	684	10	RN0312394
17	140.8	30.2	154354	2	RN470J21
18	99.2	19.1	113396	2	RN232H10
19	18.2	10.3	1184	8	AF429429
20	16.6	10.0	1399	8	PVU77939
21	15.6	9.8	36384	9	AC004178
22	15.6	9.8	104423	2	HS21UB1QU
23	15.6	9.8	112578	9	AL645927
24	15.6	9.8	114868	9	HS994E9
25	15.6	9.8	133836	9	AL662869
26	15.2	9.7	1356	8	AY052698
27	15.2	9.7	1319	8	AY057500
28	15.2	9.7	1577	8	ATHUBQ10R
29	15.2	9.7	1637	8	AY057500
30	15.2	9.7	1637	8	AY056295
31	15.2	9.7	145453	8	AC012477
32	15.2	9.7	183147	8	AC012477
33	15.2	9.7	199987	8	AC012477
34	15.2	9.7	199987	8	ATCHRIV15
35	15.2	9.7	1321	8	TOBUB14A
36	15.2	9.7	1368	8	AF429430
37	15.2	9.7	1368	8	AF429430
38	15.2	9.7	1187	8	ATU84958
39	15.2	9.7	456	8	SCFPOLY
40	15.2	9.7	4442	8	AF165330
41	15.2	9.7	156086	2	AF184280
42	15.2	9.7	179713	2	AF004126
43	15.2	9.7	3885	8	AC097297
44	15.2	9.7	7218	6	LEPUR
45	15.2	9.7	2544	8	L81141
46	15.2	9.7	871	8	ATHUBQB

#### ALIGNMENTS

RESULT 1  
AX329950  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,  
Horrikan, S., Soppet, D.R. and Weaver, Z.  
Cancer gene determination and therapeutic screening using signature  
gene sets  
Patent: WO 0194629-A 459 13-DEC-2001;  
Avalon Pharmaceuticals (US)  
Location/Qualifiers  
1: 466  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COVI 118 a 130 c 89 g 129 t

Linear PAT 09-JAN-2002

466 bp DNA

Sequence 459 from Patent WO0194629.

AX329950.1 GI:18102928



## ORIGIN

Query Match 100.0%; Score 466; DB 5; Length 466.  
 Best Local Similarity 100.0%; Pred. No. 1.6e-131; Indels 0; Gaps 0;  
 Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atcaagaacacatagagtcgggcaataatacttccatccaccatccaccacaaacttta 60  
 |||||  
 Db 1 ATCAAGAAACATAGAGTTCGGCAATATACCTTCATCCATACCCACCCAAATCITA 60  
 |||||  
 Qy 61 ctctactcatctcatctcatttaatttttgggaaatcatcagaagatatctctgaagta 120  
 |||||  
 Db 61 CTCTACTCATCTCATCTCAITTAATTTTGGGAAATCATCAGAAGATGTTGCTTGA 120  
 |||||  
 Qy 121 agagattaaagaataaagctttttgacccctcccaacaccccatccaccaggtttcac 180  
 |||||  
 Db 121 AGAGATTAAAGAATAAAGCTTTTTCACCCCTGCCAACACCCCATGCCAGGCTTCAC 180  
 |||||  
 Qy 181 cctccatacaataacatgccagaagagtaagtaccctttctgaatgccgaatctgcc 240  
 |||||  
 Db 181 CCTCCATAACATAACATGCCAGGAAGTAAGTTGCCCTTTCTGAIGCGGTAATTCGCC 240  
 |||||  
 Qy 241 atcatcttcccatcttcacgtccctttccattgcaagtcacaatctgggtctcgggat 300  
 |||||  
 Db 241 ATCATCTTCCCATCTTCCAGTCTCTTCCATTCGAAGTCACAACTCGGGTCTCAAGGAT 300  
 |||||  
 Qy 301 taccctgcttaagctcgaatctgtcttccattgacatgacacacacacacacacacacac 360  
 |||||  
 Db 301 TATACCCGCTTACGTCTGATCATGCTTTCACCTTGGCCATGAGCTGGACCTTCGCAC 360  
 |||||  
 Qy 361 ctggagagagtgctcttggctcctacacacacacacacacacacacacacacacacacac 420  
 |||||  
 Db 361 CTGGAGAGAGTGCTCTTGGCTCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 |||||  
 Qy 421 catcactggggttcac 466  
 |||||  
 Db 421 CATCAGTGGGCTTACACCACTTTCAGGGGTAAAGTGGATGGTCTT 466  
 |||||

## RESULT 2

AL645936 155874 bp DNA linear 812 31-JAN-2002  
 LOCUS Homo sapiens chromosome 6 clone XHbac-125d10, \*\*\* SEQUENCING IN  
 PROGRESS \*\*\*, in ordered pieces.  
 AL645936  
 HG: HG5\_P2; HG5\_ACTIVEFIN; HG5\_DRAFT; HG5\_FULLIOP.  
 SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (sites)  
 Leongamornlert.D.  
 Direct Submission  
 Submitted (25-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries: [clonerequests@sanger.ac.uk](mailto:clonerequests@sanger.ac.uk)  
 humqurys@sanger.ac.uk  
 On Feb 1, 2002 this sequence version replaced gi:1615238.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: [humqurys@sanger.ac.uk](mailto:humqurys@sanger.ac.uk)  
 ----- Project Information  
 Center project name: bpg126d10  
 ----- Summary Statistics  
 Assembly program: XGAP4; version 4.5  
 Sequencing vector: plasmid; 108752; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Consensus quality: 155824 bases at least Q40  
 Consensus quality: 155864 bases at least Q30  
 Consensus quality: 155872 bases at least Q20

## COMMENT

Insert size: 155874; sum-of-contigs  
 Insert size: 162470; 3.9% error; agarose-fp  
 Quality coverage: 10.66x in Q20 bases; sum-of-contigs Quality  
 coverage: 10.30x in Q20 bases; agarose-fp  
 -----  
 \* NOTE: This is a 'working draft' sequence.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 Location/Qualifiers  
 1. 155874  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /clone="XHbac-126d10"  
 /clone\_11b="CHORI-501"  
 1. 155874  
 /note="assembly-fragment:03519"  
 BASE COUNT 43601 a 36123 c 34182 g 41968 t  
 ORIGIN  
 Query Match 85.9%; Score 400.2; DB 2; Length 155874;  
 Best Local Similarity 98.5%; Pred. No. 3e-111;  
 Matches 446; Conservative 0; Mismatches 3; Indels 4; Gaps 4;  
 Qy 1 atcaagaacacatagagtcgggcaataatacttccatccaccatccaccacaaacttta 60  
 |||||  
 Db 6691 ATCAAGAAACATAGAGTTCGGCAATATACCTTCATCCATACCCACCCAAATCITA 6749  
 |||||  
 Qy 61 ctctactcatctcattcatttaatttttgggaaatcatcagaagatatctctgaagta 120  
 |||||  
 Db 6750 CTCTACTCATCTCATCTCAITTAATTTTGGGAAATCATCAGAAGATGTTGCTTGA 1209  
 |||||  
 Qy 121 agagattaaagaataaagctttttgacccctcccaacaccccatccaccaggtttcac 180  
 |||||  
 Db 6810 AGAGATTAAAGAATAAAGCTTTTTCACCCCTGCCAACACCCCATGCCAGGCTTCAC 1869  
 |||||  
 Qy 181 cctccatacaataacatgccagaagagtaagtaccctttctgaatgccgaatctgcc 240  
 |||||  
 Db 6870 TATACCCGCTTACGTCTGATCATGCTTTCACCTTGGCCATGAGCTGGACCTTCGCAC 2448  
 |||||  
 Qy 241 atcatcttcccatcttcacgtccctttccattgcaagtcacaatctgggtctcaggat 300  
 |||||  
 Db 6930 ATCACTTCCCATCTTCCAGTCTCTTCCATTCGAAGTCACAACTCGGGTCTCAGGAT 3088  
 |||||  
 Qy 361 taccctgcttaagctcgaatctgtcttccattgacatgacacacacacacacacacacac 360  
 |||||  
 Db 6989 TATACCCGCTTACGTCTGATCATGCTTTCACCTTGGCCATGAGCTGGACCTTCGCAC 7048  
 |||||  
 Qy 361 ctggagagagtgctcttggctcctacacacacacacacacacacacacacacacacacac 420  
 |||||  
 Db 7049 CTGGAGAGAGTGCTCTTGGCTCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7107  
 |||||  
 Qy 421 catcactggggttcac 453  
 |||||  
 Db 7108 CATCAGTGGGCTTACACCACTTTCAGGGGTAAAG 7139  
 |||||  
 RESULT 3  
 AF123050/2  
 LOCUS Homo sapiens ubiquitin-like protein FAT10 (FAT10) mRNA, complete  
 DEFINITION cds.  
 ACCESSION AF123050  
 VERSION AF123050.1 GI:5821714  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 971)  
 REFERENCES  
 Liu,Y.C., Pan,J., Zhang,C., Fan,W., Collinge,M., Bender,J.R. and



```
repeat_region 5815..5853
/note="MIR repeat: matches 108..146 of consensus"
repeat_region 7324..7786
/note="HCRV16 repeat: matches 596..1091 of consensus"
repeat_region 7877..8024
/note="LIM4 repeat: matches 3415..3586 of consensus"
repeat_region 8171..8565
/note="IHE1C repeat: matches 1..371 of consensus"
repeat_region 8504..8910
/note="AlusX repeat: matches 2..306 of consensus"
repeat_region 9002..9354
/note="MLT1B repeat: matches 1..390 of consensus"
repeat_region 9717..9795
/note="MIR repeat: matches 20..98 of consensus"
repeat_region 10027..10047
/note="LIMB6 repeat: matches 5767..5786 of consensus"
repeat_region 10048..10493
/note="MLT1C repeat: matches 1..466 of consensus"
repeat_region 10494..10848
/note="LIMB5 repeat: matches 5786..6145 of consensus"
repeat_region 11216..11356
/note="L2 repeat: matches 2554..2703 of consensus"
repeat_region 11386..11537
/note="MIR repeat: matches 7..167 of consensus"
repeat_region 11610..11785
/note="LIMC2 repeat: matches 4759..4931 of consensus"
repeat_region 11877..12095
/note="Alusg1 repeat: matches 1..308 of consensus"
repeat_region 12036..13044
/note="LIMC2 repeat: matches 4931..5829 of consensus"
repeat_region 13093..13176
/note="LIMB1 repeat: matches 5728..6118 of consensus"
repeat_region 13916..13991
/note="MIR repeat: matches 69..147 of consensus"
repeat_region 14700..14834
/note="MIR repeat: matches 89..226 of consensus"
repeat_region 16392..16688
/note="MER9 repeat: matches 215..511 of consensus"
repeat_region 16994..17015
/note="L1 copies 2 mer aa 100% conserved"
repeat_region 17498..17507
/note="match: GSS: Em:A0883063"
misc_feature 17524..17673
/note="match: GSS: Em:B72414"
repeat_region 17750..18033
/note="LIMC/D repeat: matches 5479..5757 of consensus"
misc_feature 18059..18968
/note="match: GSS: Em:A0743745"
repeat_region 18234..19882
/note="MER4-internal repeat: matches 2701..4900 of consensus"
repeat_region 19883..20181
/note="AluJb repeat: matches 1..292 of consensus"
repeat_region 20182..20687
/note="MER4-internal repeat: matches 2198..4701 of consensus"
repeat_region 20675..20796
/note="MER4-internal repeat: matches 2341..4464 of consensus"
repeat_region 20777..20853
/note="MER4-internal repeat: matches 1034..1114 of consensus"
repeat_region 20854..21376
/note="LIR35 repeat: matches 1..612 of consensus"
repeat_region 21377..22220
/note="MER4-internal repeat: matches 121..1337 of consensus"
repeat_region 22222..22934
/note="MER50 repeat: matches 1..711 of consensus"
repeat_region 22959..23269
/note="AluJb repeat: matches 1..307 of consensus"
repeat_region 23272..23353
/note="MER4-internal repeat: matches 1..84 of consensus"
```

```
repeat_region 23357..23776
/note="MER4C repeat: matches 1..465 of consensus"
misc_feature 25589..27225
/note="match: GSS: Em:AQ195255
match: SFS: Em:G61733"
repeat_region 28927..29156
/note="SN5 repeat: matches 1..235 of consensus"
repeat_region 30633..30690
/note="29 copies 2 mer gt 74% conserved"
repeat_region 31019..31335
/note="AluJb repeat: matches 1..311 of consensus"
repeat_region 31856..32451
/note="LIM4 repeat: matches 4650..5281 of consensus"
repeat_region 32517..32893
/note="LIM4 repeat: matches 3785..4144 of consensus"
repeat_region 32899..33125
/note="LIMEC repeat: matches 1787..2016 of consensus"
misc_feature 33093..33554
/note="match: GSS: Em:AQ620121"
repeat_region 33207..33337
/note="LIM4 repeat: matches 4708..4838 of consensus"
repeat_region 33458..33574
/note="MIR repeat: matches 144..252 of consensus"
repeat_region 33679..34067
/note="MSIA repeat: matches 1..426 of consensus"
gene complement(34169..35034)
/genes="dJ271M21.8"
CDS complement(34169..35034)
/genes="dJ271M21.8"
match: proteins: Sw:P30554 Sw:P12526 Sw:P04201 Sw:P35410
/codon_start=1
/pseudo
evidence=not_experimental
repeat_region 35631..35754
/note="L2 repeat: matches 2367..2511 of consensus"
repeat_region 36169..36279
/note="MER81 repeat: matches 3..113 of consensus"
repeat_region 36455..36829
/note="MLT1B repeat: matches 1..374 of consensus"
repeat_region 37958..38292
/note="LIMB5 repeat: matches 5817..6165 of consensus"
repeat_region 39291..39590
/note="AluSp repeat: matches 1..301 of consensus"
repeat_region 40192..41536
/note="LIM4 repeat: matches 2220..3639 of consensus"
repeat_region 41538..41700
/note="LIM4C repeat: matches 1140..1302 of consensus"
repeat_region 41701..42011
/note="AluYa5 repeat: matches 1..311 of consensus"
repeat_region 42012..43297
/note="LIM4C repeat: matches 1301..2216 of consensus"
repeat_region 43281..43540
/note="LIM5 repeat: matches 5814..6059 of consensus"
repeat_region 43538..43961
```

Query Match 85.5% Score 398.6; DB 9; Length 134292;  
Best Local Similarity 98.2% Pred. No. 9.2e-111;  
Matches 445; Conservative 0; Mismatches 4; Indels 4; Gaps 4;

Qy 1 tctaaagaaacatagagtcgggcaataatacttctaccatccacccaatctta 60  
|||||  
Db 52163 ATCAAGAAACATAGAGTT-GGGCAATACTTATCTATCCATCCCAATCTTA 52161

Qy 61 tctactcatctctcattataatttgggaatacatcaagaatgtgttcgttgagta 120  
|||||  
Db 52162 CTCIACATCACTCAITCAITTAATTTGGGAATCATCAGAGATGTGTTCGTGAGTA 52221

Qy 121 agagattaaagaaataagctttttgacccctgccaacacccatgccacggaggtcac 180  
|||||  
Db 52222 TGAGATTAAAGAAATAAGCTTTTIGACCCCTGCCAACACCCCATGCCAGGGTGGTAC 52281

Qy 191 cctcaataacaataacatgccagagagagtaagtgtgcctttctgatccgtaattctgcc 240



```

REFERENCE          2 (bases 1 to 777)
AUTHORS            Bates, E.E.M.
TITLE              Direct Submission
JOURNAL            Submitted (17-APR-1997) E.E.M. Bates, Schering-Plough, Laboratory
                  for Immunological Research, 27 Chemin des Peupliers, BP11, 95571
                  Bardilly Cedex, FRANCE
FEATURES
  Source           1..777
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /cell_type="dendritic cells"
                  /codon_start=1
                  /product="diubiquitin"
                  /protein_id="CAA73200.1"
                  /db_xref="GI:2546964"
                  /db_xref="SPTREMBL:O15205"
                  /translation="MAPNASCILCVHVRSEENDLMTFDANPYDSVKIKHEVSKIKVP
                  VQDVLKLSKILKPRSLSSGIDKEKIIHLTKVKFSDEELPLFLVLESDEAKRH
                  LQVRESSVAOVKAMIEIKTGIPIETQIVTCNGKRLDEKRWADYGIKRNLLFLAS
                  YCIGG"
  polyA_signal     223 a 153 c 195 g 206 t
                  748..753
ORIGIN
Query Match      85.2%; Score 397; DB 9; Length 777;
Best Local Similarity 98.0%; Pred. No. 2.7e-110;
Matches 444; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

OY 1 atcaagaacatagattcgaggcaataatacttcattcatccaccatcccccacaaatctta 60
DB 591 ATCAAGAACAATAGAGTII-GGGCAATATATCTTCACTACCCATCCCAATCTTA 633
OY 61 ctctactcatctcattctcatttaattttgggaataatcatcagaagatgtttcgttgagta 120
DB 632 CTCTACTCATCTCATCTCAITTAATTTTGGGAATATCATCAGAAGATGTTTCGTIGAGTA 573
OY 121 agagattaaagaataagctttttgacctgtgcaacaccccatgccaggtggtcac 180
DB 572 AGAGATTAAAGAATAAGCTTTTIGACCCCTGCCACACCCCATCCCAAGSGTGGTCAC 513
OY 181 ctctccatacataatgacgaggagaagtagtgccctttctgagtcgcaataatctgcc 240
DB 512 CTCTCAATACATATAGATGCCAGGAAGTAGTAAGTTGCCCTTTCATGCCGTAATCTGCC 453
OY 241 atcattctccatcttcacgtctcatttccatttgcagagtcacaatctgggtctcaggat 300
DB 452 ATCATCTTCCCATCTTCCAGTCT-CITTTCAITGCAAGTCACAATCTGGGTCTCAGGGAT 334
OY 301 tataccgctttagtctgacattgtcttccattgtgccactgaagctgggaccttcgcac 360
DB 393 TATACCCGCTTAGTCTCGATCATTGCTTTCACATTGTCGCACCTGAGCTGGACCTTCGCAC 334
OY 361 ctggagaggagtgctctttgacctcattcactgactccacagaagaagaagagagctcct 420
DB 333 CT-GGAGGAGGTGCTCTTIGCCICATCACCCTGACTCCACAGAAGAACAAGSGAGCTCCT 275
OY 421 catcactggggttccaccacttccagggttaag 453
DB 274 CATCACT-GGGCTTCACCACATTCAGGGTAAGG 243

RESULT 7
AC004179
LOCUS              Homo sapiens chromosome 6 clone UWGC:CB map p21. *** SEQUENCING IN
DEFINITION         AC004179 41069 bp DNA linear HTG 23-FEB-1998
PROGRESS ***; 1 ordered pieces.
ACCESSION          AC004179
VERSION            AC004179.1 GI:2905981
KEYWORDS           HTG; HTGS_PHASE2.
SOURCE             human.
ORGANISM           Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 41069)
JANER M., Guillaudoux, I., Vu, O., Kutyavin, T., Harter, H. and
Geraghty, D.E.
Large scale sequence analysis of the human MHC class I region
Unpublished (1998)
Fred Hutchinson Cancer Research Center
The Clinical Research Division
1100 Fairview Ave. N., P.O. Box 19024
Seattle, WA 98109-1024

2 (bases 1 to 41069)
Geraghty, D.E. and Olson, M.V.
Direct Submission
Submitted (23-FEB-1998) Human Genome Center, University of
Washington, Box 352145, Seattle, WA 98195, USA
University of Washington Human Genome Center
Box 352145 Seattle, WA 98195
Contact: Daniel E. Geraghty (geraghty@fhcrc.org).
* NOTE: This is a 'working draft' sequence. It currently
  * consists of 1 contigs. Gaps between the contigs
  * are represented as runs of N. The order of the pieces
  * is believed to be correct as given, however the sizes
  * of the gaps between them are based on estimates that have
  * been provided by the submitter.
  * This sequence will be replaced
  * by the finished sequence as soon as it is available and
  * the accession number will be preserved.
  * 1 41069: contig of 41069 bp in length.
FEATURES
  Location/Qualifiers
    1..41069
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /chromosome="6"
      /map="p21"
      /clone="UWGC:CB"
BASE COUNT 12023 a 9129 c 8744 g 11173 t
ORIGIN

Query Match      85.2%; Score 397; DB 2; Length 41069;
Best Local Similarity 98.0%; Pred. No. 2.7e-110;
Matches 444; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

OY 1 atcaagaacatagattcgaggcaataatacttcattcatccaccatcccccacaaatctta 60
DB 18419 ATCAAGAACAATAGAGTII-GGGCAATATATCTTCACTACCCATCCCAATCTTA 18477
OY 61 ctctactcatctcattctcatttaattttgggaataatcatcagaagatgtttcgttgagta 120
DB 18478 CTCTACTCATCTCATCTCAITTAATTTTGGGAATATCATCAGAAGATGTTTCGTIGAGTA 18537
OY 121 agagattaaagaataagctttttgacctgtgcaacaccccatgccaggtggtcac 180
DB 18538 AGAGATTAAAGAATAAGCTTTTIGACCCCTGCCACACCCCATCCCAAGSGTGGTCAC 18597
OY 181 ctctccatacataatgacgaggagaagtagtgccctttctgagtcgcaataatctgcc 240
DB 18539 CTCTCAATACATATAGATGCCAGGAAGTAGTAAGTTGCCCTTTCATGCCGTAATCTGCC 18657
OY 241 atcattctccatcttcacgtctcatttccattgtgccactgaagctgggtctcaggat 300
DB 18659 ATCATCTTCCCATCTTCCAGTCT-CITTTCAITGCAAGTCACAATCTGGGTCTCAGGGAT 18716
OY 301 tataccgctttagtctgacattgtcttccattgtgccactgaagctggaccttcgcac 360
DB 18717 TATACCCGCTTAGTCTCGATCATTGCTTTCACATTGTCGCACCTGAGCTGGACCTTCGCAC 18776
OY 361 ctggagaggagtgctctttgacctcattcactcactccacagaagaagaagagagctcct 420
DB 18777 CT-GGAGGAGGTGCTCTTIGCCICATCACCCTGACTCCACAGAAGAACAAGSGAGCTCCT 18835

```

```

O> 421 caccactgggcttcaccactttccaggagtaag 453
||||| ||||||| ||||||| ||||||| |||
Db 18836 CATCATTGGGCTTACACACATTCAGGGTAAGG 18867

RESULT 8
AL571618/c 129567 bp DNA linear HIG 01-FEB-2002
LOCUS Homo sapiens chromosome 6 clone XHbac-140P7, ** SEQUENCING IN
DEFINITION PROGRESS **, 2 unordered pieces.
ACCESSION AL571618
VERSION AL571618.2 GI:18491494
KEYWORDS HIG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HIGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Direct Submission
Submitted (31-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Feb 4, 2002 this sequence version replaced gi:18477174.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BRGL40P7
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752: 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 129267 bases at least Q40
Consensus quality: 129282 bases at least Q30
Consensus quality: 129363 bases at least Q20
Insert size: 129467; sum-of-contigs
Insert size: 135304; 5.8% error; agarose-fp
Quality coverage: 10.92x in Q20 bases; sum-of-contigs Quality
coverage: 10.45x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 103125: contig of 103125 bp in length
* 103126 103225: gap of 100 bp
* 103226 129567: contig of 26342 bp in length.
Location/Qualifiers
1..129567
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="XHbac-140P7"
/clone_lib="CHORI-501"
1..103125
/notes="assembly fragment:02311
fragment_chain:1
clone_end:17
vector_side:left"
103226..129567
/notes="assembly fragment:00372
fragment_chain:1
clone_end:SP6
vector_side:right"

misc_feature 36114 a 27532 c 28517 g 37304 t 100 others

misc_feature
FEATURES
Source

```

## ORIGIN

```

Query: Watch 85.2% Score 397; DB 2: Length 129567;
Best Local Similarity 98.0%; Pred. No. 2.8e-110;
Matches 444; Conservative 0; Mismatches 5; Indels 4; Gaps 4:

O> 1 atcaagaagaacatagattcgggcaataatacttctaccctcccccccaaatctta 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 4382 ATCAAGAAGAAATAGAGTT-GGGCAATATATCTTCACTCTACCCATCCCAAAATCTTA 4324
O> 61 ctctactcatctcatctcatttaatttgggaataatcatcagaagatgtgttcgttaagta 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 4323 CTCTACTCATCTCATCTCTCAITTAATTTTGGAAATCATCAGAACATGTGTCGTTCAGTA 4264
O> 121 agagattaaaagaataagctttttgacccctcccaaccccccatgccagggtgtgtac 180
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 4263 AGAGATTAAAGAAATAAGCTTTTIGACCCCTGCCAACCCCATCCCGAGGTGGTCA 4204
O> 181 cctccataacaataacatgccaggaagaagtaagtgcctttctgataccgtaactctgc 240
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 4203 CTTCCAAATACAAATGAATGCCAGGAGTAAGTTGCCCCCTTCTGATGCCGTAACTGCC 4144
O> 241 atcatcttcccatcttccagctctctcttccattgcaagtcacaaatctgggtctcaggat 300
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 4143 AICATCTTCCATCTTCCAGTCT-CTTTCCATTGCAAGTCACAATCTGGGTCTCAGGGAT 4085
O> 301 tataccgctttagctcgcattgtcttccattgtgcactgaactgaactctgcac 360
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 4084 TATACCCGCTTAGTCTCGATCATGTCTTTCACCTTGCCACTGAGCTGACCTTCGCAC 4025
O> 361 ctggagagagtgctctcttgcctcatcactgactcacaacaaagagggagctct 420
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 4024 CT-GGAGGAGTGCGCTTTTGGCTTCATCACCCTGACTCCACAAGAAACAAGGGAGCTCT 3956
O> 421 catcattgggcttcacacactttcaggagtaag 453
||||| ||||||| ||||||| |||
Db 3965 CATCACT-GGGCTTACACCACTTTCAGGTAAGG 3934

RESULT 9
AC006137/c 129806 bp DNA linear PRI 19-OCT-2001
LOCUS Homo sapiens clone SCB-254N2 (UMGC:rg254N02) from 6p21, complete
DEFINITION sequence.
ACCESSION AC006137
VERSION AC006137.3 GI:16271961
KEYWORDS HIG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 129806)
Janer, M., Guillaudoux, T., Vu, Q., Kutayavin, T., Harter, H. and
Geraghty, D.E.
Large scale sequence analysis of the human MHC class I region
Unpublished
Fred Hutchinson Cancer Research Center
The Clinical Research Division
1100 Fairview Ave. N., P.O. Box 19024
Seattle, WA 98109-1024
2 (bases 1 to 129806)
Geraghty, D.E. and Olson, M.V.
Direct Submission
Submitted (05-DEC-1998) Human Genome Center, University of
Washington, Box 352145, Seattle, WA 98195, USA
University of Washington Human Genome Center
Box 352145, Seattle, WA 98195
Contact: Daniel E. Geraghty (geraghty@fhcrc.org)
3 (bases 1 to 129806)
Kaul, R.K. and Haugen, E.D.
Direct Submission
Submitted (24-AUG-2000) Genome Center, University of Washington.
JOURNAL

```



KEYWORDS MGC.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
TITLE 1 (bases 1 to 791)  
JOURNAL Strausberg, P.  
Direct Submission  
Submitted (15-AUG-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cdaps-r@mail.nih.gov](mailto:cdaps-r@mail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [villalob@bcm.tmc.edu](mailto:villalob@bcm.tmc.edu)  
Villalob, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,  
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,  
Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 28 Row: 1 Column: 19  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 5821714.

FEATURES  
source Location/Qualifiers  
1..791  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="MGC:21200 IMAGE:4517233"  
/tissue\_type="Bladder, transitional cell papilloma"  
/clone\_lib="NH\_MGC\_93"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT5"  
25..522  
/codon\_start=1  
/product="Similar to diubiquitin"  
/protein\_id="AAH12472.1"  
/db\_xref="GI:15214580"  
/translation="MAPNASCLVHVRSEENDLMTFDANPYSVFKIKHYRSKTKVF  
VDDVLLGSKILPRSSSYGIDKERTIHLTKVWPSEELFLFVSGDEAKPH  
LLQVRRSSVAQVKAMIEIKTGIIIPETQIVTCNGKRLDGRMHADYGIKGNLLFLAC  
YCIGG"  
BASE COUNT 239 a 153 c 197 g 202 t  
ORIGIN

Query Match 84.8%; Score 395.4; DB 9; Length 791;  
Best Local Similarity 97.8%; Pred. No. 6.5e-110;  
Matches 443; Conservative 0; Mismatches 5; Indels 4; Gaps 4;  
CQ 1 atcaagaacatagagttcggacaataatacttctcctaccatcccccaaaactta 60  
DB 697 ATCAAGAACAATAGAGTT-GGGCAATATACITTCCTACCCATCCAGCCCAACICTTA 539  
CQ 61 ctctactcatctctcattcattatttgggaaatcatcagaagatgtgttcgttagta 120  
DB 638 CTCTACTCATCTCATCTCATTATTTTGGGAAATCATCAGAGATGTGTCTGTCAGTA 579  
CQ 121 agagattaaagaataagctttttgacccctcccaacaccccatgccagggtgtcac 180  
DB 578 AGAGATTAAAGGAAATTAAGCTTTTTCACCCCTGCCAACACCCCATGCCAGGGGTGCAC 519  
CQ 181 ctcccaataacaataacatgccaggagatgaatttccctttctcattgacgtgacgttcac 240  
DB 518 CTCCCAATGAATAACATGCCAGGAAGATGAAGTTGCCCTTCTGATGCCGTAATTCGCC 459

CQ 241 atcatcttcccatcttccagttctcttccattgcaagtccacaatctgggtctcagggaat 300  
DB 458 ATCACTTCCCATCTTCCAGTCT-CTTCCATTCAGTACCAATCTGGGTCTCAGGAT 400  
CQ 301 tataccctcttagctctcgatcatcttgcatttgcatttgccttgccactgagctgaccttcgac 360  
DB 399 TATACCCGTCTTAGTCTCGATCATGTCTTCACTTGTGCCACTGAGCTGGACCTTCGCAC 340  
CQ 361 ctgagagaagatgcctcttccctcatcaccctgactcacaagaacaaagggcagctcct 420  
DB 399 CT-GGAGGAGTGCCTCTTTGCCCTCATCACTGACTCCACAAGAAACAAGGCGAGCTCT 281  
CQ 421 catcactggggcttcaccactttcaggggtaag 453  
DB 250 CATCACI-GGGCTTCACCACTTTTCAGGGTAAAG 249  
RESULT 11  
AX237229/c  
LOCUS AX237229  
DEFINITION Sequence 205 from Patent WO0164886.  
ACCESSION AX237229  
VERSION AX237229.1 GI:15796783  
KEYWORDS human.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 474)  
AUTHORS Gaiger, A., Algate, P.A. and Mannion, J.  
TITLE Compositions and methods for the detection, diagnosis and therapy  
of hematological malignancies  
JOURNAL Patent: WO 0164886-A 205 07-SEP-2001;  
CORIXA CORPORATION (US)  
FEATURES  
source Location/Qualifiers  
1..474  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 165 a 70 c 111 g 126 t  
ORIGIN

Query Match 71.8%; Score 334.4; DB 6; Length 474;  
Best Local Similarity 99.2%; Pred. No. 2.9e-91;  
Matches 357; Conservative 0; Mismatches 1; Indels 2; Gaps 2;  
CQ 1 atcaagaacatagagttcggacaataatacttctcctaccatcccccaaaactta 60  
DB 358 ATCAAGAACAATAGAGTT-GGGCAATATACITTCCTACCCATCCCAACICTTA 300  
CQ 61 ctctactcatctctcattcattatttgggaaatcatcagaagatgtgttcgttagta 120  
DB 299 CTCTACTCATCTCATCTCATTATTTTGGGAAATCATCAGAGATGTGTCTGTCAGTA 240  
CQ 121 agagattaaagaataaactttttgacccctcccaacaccccatgccagggtgtcac 180  
DB 239 AGAGATTAAAGAAATTAAGCTTTTTCCTGCCAACACCCCATGCCAGGGGTGCAC 180  
CQ 181 ctcccaataacaataacatgccaggagagtaattgcccctttctctgacgttaacttcgac 240  
DB 179 CTCCCAATGAATAACATGCCAGGAAGATGAAGTTGCCCTTCTGATGCCGTAATTCGCC 120  
CQ 241 atcatcttcccatcttccagttctcctttccattgcaagtcacaatctgggtctcagggaat 300  
DB 119 ATCACTTCCCATCTTCCAGTCT-CTTCCATTCGAAGTCACAAATCTGGGTCTCAGGAT 61  
CQ 201 ataccctgttagctcgcattcatttcccttctcattgacgtgacgtgacgttcacac 360  
DB 60 TATACCCGTCTTAGTCTCGATCACTTCTTCACTTGTGCCACITGAGCTGGACCTTCGCAC 1

RESULT 12



AX237670/c	AX237670	474 bp	DNA	linear	PAT 25-SEP-2001
LOCUS	Sequence 545 from Patent WO0164885.				
DEFINITION	AX237670				
ACCESSION	AX237670				
VERSION	AX237670.1	GI:15797224			
KEYWORDS	human.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 474)				
TITLE	Gaiger, A., Algate, P.A. and Mannion, J.				
JOURNAL	Compositions and methods for the detection, diagnosis and therapy of hematological malignancies				
FEATURES	Patent: WO 0164885-A 546 07-SEP-2001;				
source	CORIXA CORPORATION (US)				
BASE COUNT	Location/Qualifiers				
ORIGIN	1..474				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	165 a 70 c 111 g 125 t 2 others				
Query Match	71.8%; Score 334.4; DB 6; Length 174;				
Best Local Similarity	99.2%; Pred. No. 2.9e-91;				
Matches 357:	Conservative 0; Mismatches 1; Indels 2; Gaps 2;				
QY	1 atcaagaacatagattcgggcaataatacttctaccctaccacccaccccaaatctta 60				
Db	358 ATCAAGAACAATAGAGTII-GGGCAATATATCTTCACTACCCATCCACCGAACTTAA 300				
QY	61 ctctact 120				
Db	299 CTCTACT 240				
QY	121 agagattaaagaataaagctttttgacccctcgccacaccccatgccacgggtggtcac 180				
Db	239 AGAGATTAAAGAATAAAGCTTTTGTATCCCTGCCACACCCCATATGCCAGGTCTGTCAC 192				
QY	181 cctccaatcaataaacatgccagaagagtaagtggcctttctgattgacctaatctgcc 240				
Db	179 CTCCCAATCAATAAATGATGCCAGGAAGAGTAAGTGGCTTTCTGATGCCGTAATGTGC 120				
QY	241 atcatctcccatcttccagctctcttctccatcgcaagtcaaatctgggtcagggat 300				
Db	119 ATCATCTCCCATCTTCCAGTCT-CTTTCATTCGAAGTCACAACTGGGTCTAGGGAT 61				
QY	301 tataccgctcttagctcgatcattctcttctctctctctctctctctctctctctctct 360				
Db	50 TATACCGCTCTTAGCTCGATCAATGCTTTCAGTCTGTGCCACTGAGCTGGACCTTGGCAC 1				
RESULT 13					
LOCUS	AX237272/c	453 bp	DNA	linear	PAT 26-SEP-2001
DEFINITION	Sequence 248 from Patent WO0164885.				
ACCESSION	AX237272				
VERSION	AX237272.1	GI:15796826			
KEYWORDS	human.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 453)				
TITLE	Gaiger, A., Algate, P.A. and Mannion, J.				
JOURNAL	Compositions and methods for the detection, diagnosis and therapy of hematological malignancies				
FEATURES	Patent: WO 0164885-A 248 07-SEP-2001;				
source	CORIXA CORPORATION (US)				
	Location/Qualifiers				
	1..453,				
	/organism="Homo sapiens"				





Result No.	Query Match	Score	Length	DB	ID	Description
C 1	85.8	399.8	833	21	AAC78081	Human cancer assoc
C 2	85.2	397	777	19	AAV36265	Human diubiquitin
C 3	71.8	334.4	474	22	AAK54480	Human haematologic
C 4	71.8	334.4	474	22	AAK54921	Human haematologic
C 5	69.4	323.4	453	22	AAK54523	Human haematologic
C 6	50.3	234.5	566	23	AAK568568	DNA encoding novel
C 7	50.3	234.6	503	1129	AAK590737	DNA encoding novel
C 8	49.3	229.8	684	23	AAK71154	DNA encoding novel
C 9	49.3	229.8	684	23	AAK71927	DNA encoding novel

Human: cancer associated gene; cancer antigen; detection; cancer; diagnosis; cytostatic; proliferative; vulnerable; immunomodulator; antidiabetic; antiasthmatic; antithrombotic; antiarthritic; antiviral; antiinflammatory; antithyroid; antiangiogenic; antibacterial; cardiant; dermatological; neuroprotective; thrombolytic; coagulant; neurotropic; vasotonic; antipsoriatic; antiangiogenic; gene therapy; inflammation; immune disorder; haematopoietic cell disorder; autoimmune disorder; allergic reaction; graft versus host disease; organ rejection; haemostatic; thrombolytic; cardiovascular disorder; infection; neurological disease; drug screening; ss.

Hetero sapiens

WO20055350-A1

21-SEP-2000.

05-MAR-2000: 2000WO-US05882.

12-WAF-1999; 99US-0124270.

(HUMA-) HUMAN GENOME SCI INC.





```

PR 01-MAY-2000: 2000US-0200999.
PR 04-MAY-2000: 2000US-0202084.
PR 22-MAY-2000: 2000US-0206201.
PR 14-JUL-2000: 2000US-0218950.
PR 03-AUG-2000: 2000US-0222903.
PR 04-AUG-2000: 2000US-0223416.
PR 07-AUG-2000: 2000US-0223378.
XX
XX (CORI-) CORIXA CORP.
PA
XX Gaiger A, Algate PA, Mannion J;
XX WPI: 2001-514842/56.
XX
XX Compositions and methods for the detection of hematological
XX malignancies, e.g. chronic lymphocytic leukemia, lymphoma, follicular
XX lymphoma and Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX
XX Claim 31: Page 499; 1252pp; English.
XX
XX The present invention relates to compositions and methods for the
XX detection, diagnosis and therapy of hematological malignancies. The
XX present sequence is the coding sequence of a human hematological
XX malignancy related antigen. The methods of the present invention comprise
XX detecting the presence of hematological malignancy related antigen(s) in
XX a sample obtained from the patient (an increased level of the
XX polypeptide, compared to an unaffected individual, is indicative of an
XX increased risk). Hematological malignancies which can be treated using
XX the present invention are chronic lymphocytic leukaemia, lymphoma,
XX follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
XX cell non-Hodgkin's lymphoma.
XX
XX Sequence 474 BP: 155 A: 70 C: 111 G: 126 T: 2 other:

Query Match      71.8%; Score 334.4; DB 22; Length 474;
Best Local Similarity 99.2%; Pred. No. 1.6e-98;
Matches 357; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1 atcaagaacacatagagttcgggcaataatactctaccatccaccacccaaatctta 60
DB 358 ATCAAGAACAACATAGAGTT-GGGCAATATACITTCATCCATCCCATCCCAATCTTA 300
QY 61 ctctactcatctctctcatttaatttgggaaatcatcagaagatgtgtcgttgagta 120
DB 299 CTCTACTCATCTCAITTCATTAATTTTGGGAAATCATCAGAGATGTGTCTGTGAGTA 240
QY 121 agagattaaaagaataaagctttttgacccctgtcccaacaccccatgccagggtgtcac 180
DB 239 AGAGATTAAAGAAATAAGCTTTTTCATCCCTGCCAACACCCCATGCCAGGGTGTGCAC 180
QY 181 cctccaatacaataacatccagagaagataagttgcccctttctgtatccgtaactcac 240
DB 179 CTCTCAATACATATACATGCCAGAGAGTAAGTTCCTTTCTGTATGCCGTAAATCTGCC 120
QY 241 atcatcttcccatcttccagctcccttccattgcaagtcacaaatctgggtctcagggat 300
DB 119 ATCATCTTCCCATCTTCCAGTCT-CITTTCCATTCGAAGTCACAAATCTGGGTCTCAGGAT 61
QY 301 tataccgctttagctcgatcttgcctttcacttgcacactgagctggagcccttcgac 360
DB 60 TATACCCGCTTATAGTCGATCATTTGCTTTTACCTTGTGCCACTGAGCTGACCTTCGCAC 1

RESULT 5
AAK54523/c
ID AAK54523 standard; cDNA: 453 BP.
XX
XX AAK54523:
XX
XX 13-NOV-2001 (first entry)
XX
XX Human hematological malignancy-related antigen coding sequence #248.

```

```

XX
XX Human: cytostatic; vascular; gene therapy; vaccine; lymphoma;
XX hematological malignancy; antigen; chronic lymphocytic leukaemia;
XX follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX
XX Hcto sapiens.
XX
XX WO200154895-A2.
XX
XX 07-SEP-2001.
XX
XX 01-MAR-2001; 2001WO-US07272.
XX
XX 01-MAR-2000: 2000US-0186126.
XX 17-MAR-2000: 2000US-0190479.
XX 27-APR-2000: 2000US-0200545.
XX 23-APR-2000: 2000US-0200303.
XX 29-APR-2000: 2000US-0200779.
XX 01-MAY-2000: 2000US-0200999.
XX 04-MAY-2000: 2000US-0202084.
XX 22-MAY-2000: 2000US-0206201.
XX 14-JUL-2000: 2000US-0218950.
XX 03-AUG-2000: 2000US-0222903.
XX 04-AUG-2000: 2000US-0223416.
XX 07-AUG-2000: 2000US-0223378.
XX
XX (CORI-) CORIXA CORP.
PA
XX Gaiger A, Algate PA, Mannion J;
XX WPI: 2001-514842/56.
XX
XX Compositions and methods for the detection of hematological
XX malignancies, e.g. chronic lymphocytic leukemia, lymphoma, follicular
XX lymphoma and Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX
XX Claim 31: Page 399; 1252pp; English.
XX
XX The present invention relates to compositions and methods for the
XX detection, diagnosis and therapy of hematological malignancies. The
XX present sequence is the coding sequence of a human hematological
XX malignancy related antigen. The methods of the present invention comprise
XX detecting the presence of hematological malignancy related antigen(s) in
XX a sample obtained from the patient (an increased level of the
XX polypeptide, compared to an unaffected individual, is indicative of an
XX increased risk). Hematological malignancies which can be treated using
XX the present invention are chronic lymphocytic leukaemia, lymphoma,
XX follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
XX cell non-Hodgkin's lymphoma.
XX
XX Sequence 453 BP: 159 A: 67 C: 103 G: 123 T: 1 other:

Query Match      69.4%; Score 323.4; DB 22; Length 453;
Best Local Similarity 99.1%; Pred. No. 6.5e-95;
Matches 346; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1 atcaagaacacatagagttcgggcaataatactctaccatccaccacccaaatctta 60
DB 347 ATCAAGAACAACATAGAGTT-GGGCAATATACITTCATCCATCCCATCCCAATCTTA 289
QY 51 ctctactcatctctcattcatttaatttgggaaatcatcagaagatgtgtcgttgagta 120
DB 299 CTCTACTCATCTCAITTCATTAATTTTGGGAAATCATCAGAGATGTGTCTGTGAGTA 229
QY 121 agagattaaaagaataaagctttttgacccctgtcccaacaccccatgccagggtgtcac 180
DB 228 AGAGATTAAAGAAATAAGCTTTTTCATCCCTGCCAACACCCCATGCCAGGGTGTGCAC 169
QY 161 cctccaatacaataacatccagagaagataagttgcccctttctgtatccgtaactcac 240
DB 159 CTCTCAATACATATACATGCCAGGAGAGTAAGTTCCTTTCTGTATGCCGTAAATCTGCC 109

```







PF 30-MAR-2001: 2001WO-US08631.  
XX  
PP 31-MAR-2000: 2000US-0540217.  
PR 23-AUG-2000: 2000US-0549157.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Drmanac RT, Liu C, Tang YI:  
XX  
XX WPI: 2001-639362/73.  
XX P-PSDB: ABG07740.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity  
XX  
XX Claim 1: SEQ ID No 7731: 103pp: English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX and gene mapping, and in recombinant production of (II). The  
XX polynucleotides are also used in diagnostics as expressed sequence tags  
XX for identifying expressed genes. (I) is useful in gene therapy techniques  
XX to restore normal activity of (II) or to treat disease states involving  
XX (II). (II) is useful for generating antibodies against it, detecting or  
XX quantitating a polypeptide in tissue, as molecular weight markers and as  
XX a food supplement. (II) and its binding partners are useful in medical  
XX imaging of sites expressing (II). (I) and (II) are useful for treating  
XX disorders involving aberrant protein expression or biological activity.  
XX The polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. AAS64197-AAS94564 represent novel human  
XX diagnostic coding sequences of the invention.  
XX Note: The sequence data for this patent did not appear in the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 684 BP: 233 A: 121 C: 200 G: 130 T: 0 other;  
XX  
XX Query Match 49.3%; Score 229.8; DB 23: Length 684;  
XX Best Local Similarity 96.4%; Pred. No. 2.1e-64;  
XX Matches 267; Conservative 0; Mismatches 7; Indels 3; Gaps 3;  
OY 177 tcaccctccaataacaatgccaggagagtaagttgccccttctgataccgtaatc 236  
DB 684 TCACCTCCCAATAAATAAGATGCCAGGAAGAGTAAGTTGCCCTTCTGATGCCGTAATC 625  
OY 237 tgccatcatcttcccattctccagctctcttccattgcaagtccaactctgggtctcag 296  
DB 624 TGCCATCACTTCCCACTTCCAGTCT-CTTCCATTCGAAGTCACAACTCGGGTCTCAG 566  
OY 297 gattataccctgcttagttctgatactgcttcttccattgcaagtccaactctgggtctcag 356  
DB 566 GGATTATACCGCTTAGTTCGATCATCTTCTTACCTTGTGCGCACTGAGCTGGACCTTC 506  
OY 357 gcaactggggaggagtgctcttctgctcatcatcactgactccaagaagaaggaagc 416  
DB 505 GCACCT-GGAGGAGGTGCTCTTCTTCCCTCATCACTTGGCTGCCAAGAAAGGGGACG 447  
OY 417 tctcatcactggggtctcaccacttccagggttaag 453  
DB 446 TCTCATCACT-GGGCTTACCACTTTCAGGGTAAG 411  
RESULT 10  
AAV36266/c  
ID AAV36266 standard: cDNA; 496 BP.  
XX

AC AAV36266:  
XX  
XX 29-SEP-1998 (first entry)  
XX  
XX Mouse diubiquitin A05F12 cDNA.  
XX  
XX Diubiquitin: mouse; dendritic cell; DC protein; cancer; diagnosis;  
XX therapy: A05F12; ss.  
XX  
XX Mus sp.  
XX  
XX Key: Location/Qualifiers  
XX CDS 8..495  
XX /\*tag= a  
XX  
XX W09823747-A2.  
XX  
XX 01-JUN-1998.  
XX  
XX 25-NOV-1997: 97WO-US20811.  
XX  
XX 11-DEC-1996: 96US-0763455.  
XX 27-NOV-1996: 96US-0031806.  
XX  
XX (SCHE ) SCHERING CORP.  
XX  
XX Bar-Chateau J, Bates EEM, Caux C, De Saint-Vis BM;  
XX Lecocque SJE:  
XX  
XX WFI: 3298-322727/28.  
XX P-FSD: AAW60832.  
XX  
XX Dendritic cell genes - useful for developing products for diagnosis  
XX and treatment of e.g. cancerous or degenerative conditions  
XX  
XX Example: Page 75; 92pp: English.  
XX  
XX This cDNA clone codes for a mouse diubiquitin protein (see AAW50832).  
XX designated A05F12, which contains 2 ubiquitin domains and which  
XX exhibits similarity to the monoclonal non-specific suppressor  
XX factor beta produced by mouse and human T cells. The clone can  
XX be expressed in a wide variety of host cells for the synthesis  
XX of full-length A05F12 or polypeptide fragments. The novel DC  
XX proteins A05F12 (see AAW50831-32), A07C03 (1g family) (see AAW50833-35),  
XX and E2B02 (lysosome associated membrane protein (LAMP)-like) (see  
XX AAW50935) are found in activated dendritic cells. The proteins play  
XX a role in the regulation or development of haematopoietic cells,  
XX e.g. lymphoid cells, which affect immunological responses, e.g.  
XX antigen presentation and the resulting effector functions. The  
XX novel nucleic acids, polypeptides, and antibodies raised against  
XX them, can be used in treatment of conditions associated with  
XX abnormal physiology or development, including abnormal  
XX proliferation, e.g. cancerous or degenerative conditions. The  
XX products can also be used for detection, diagnosis or drug  
XX screening.  
XX  
XX Sequence 496 BP: 147 A: 115 C: 125 G: 109 T: 0 other;  
XX  
XX Query Match 30.5%; Score 142; DB 19: Length 496;  
XX Best Local Similarity 74.8%; Pred. No. 6.5e-36;  
XX Matches 217; Conservative 0; Mismatches 70; Indels 3; Gaps 3;  
OY 177 tcaccctccaataacaatgccaggagagtaagttgccccttctgataccgtaatc 236  
DB 496 TCATCCCCAGTGCAGTGTGTGTTCAGAAAGCAAACTGCCACTCTTGAITGTGTGATC 437  
OY 237 tgccatcatcttcccattctccagctctcttccattgcaagtccaactctgggtctcag 296  
DB 496 AGCATGATCTTTCATCTCCAG-CTTCTTCCGTTCGAATTCACAAGCTGCTCTTAC 378  
OY 297 gattataccctgcttagttctgatactgcttcttccattgcaagtccaactctgggtctcag 356  
XX

Db 377 GGATCAGAGGTCACACTTCGATCACTCTTCCACCTGGGCGCACACGATGGATCTTC 318  
 Oy 357 gcacctggaggaggtgctcttgcctccatccactgactccacaagaacaaggcagc 416  
 Db 317 G-AACTCGGAGGAGGCTTTGGCCCTCGTTTGGACATCCACAGAAACAGGGCAGC 259  
 Oy 417 tccctcatcactgggcttcaccatttcagggttaagtggtggtctt 465  
 Db 258 TCTTCACTACT-GGGCTTCACCACTTCAGGGTAAGGTGGATAGTGGTTT 210

## RESULT 11

AAL33488

ID AAL33488 standard: DNA: 51 BP.

XX

AC AAL33488;

XX

XX 24-JAN-2002 (first entry);

XX Human SNP oligonucleotide #6696.

XX

XX Immunosuppressive: immunostimulatory; antiinflammatory; cytostatic;  
 KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;  
 KW amyloid protein; angiotensin; apoptosis related protein; cadherin;  
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
 KW complement related protein; cytochrome; kinase; cytokine; interferon;  
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
 KW multifactorial disease; autoimmune disease; infection;  
 KW nervous system disease; ss.

XX Homo sapiens.

OS

XX WO200147944-A2.

XX PD 05-JUL-2001.

XX PF 28-DEC-2000: 2000WO-US35498.

XX PR 28-DEC-1999: 99US-0173419.

XX PR 27-DEC-2000: 2000US-0173419.

XX PA (CURA-) CURAGEN CORP.

XX PI Shimkets RA, Leach M;

XX DR WPI; 2001-465210/50.

XX PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
 PT oncogenes and histones, useful for diagnosing and treating, e.g.  
 PT cancer, autoimmune diseases and infections

XX PS Claim 1: Page 3293; 4143pp; English.

XX CC The present invention relates to oligonucleotides encoding polymorphic  
 CC variants of proteins related to amylases, amyloid proteins, angiotensin,  
 CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,  
 CC histones, kinases, colony stimulating factors, complement related  
 CC proteins, cytochromes, kinases, cytokines, interferons, interleukins,  
 CC G-protein coupled receptors and thioesterases. The present sequence is  
 CC one such oligonucleotide. The oligonucleotides and the peptides encoded  
 CC by them may be used in the prevention, diagnosis and the treatment of  
 CC diseases associated with inappropriate expression of the proteins listed  
 CC above. Disorders that may be prevented, diagnosed and/or treated include  
 CC multifactorial diseases with a genetic component, such as autoimmune  
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,  
 CC systemic lupus erythematosus and Grave's disease), inflammation, cancer  
 CC (e.g. cancers of the bladder, brain, breast, colon and kidney, cancer  
 CC leukaemia), diseases of the nervous system and an infection of pathogenic  
 CC organisms.

XX S0 Sequence 51 BP: 15 A; 13 C; 12 G; 11 T: 0 other;

Query Match 10.9%; Score 51; DB 22; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 8.6e-07;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 171 ggggtggtcactccctcccaataacataatgcccaggaagaagtaagtgcctt 221  
 Db 1 ggggtggtcactccctcccaataacataatgcccaggaagaagtaagtgcctt 51

## RESULT 12

AAC42164/c

ID AAC42164 standard: DNA: 934 BP.

XX

XX AAC42164;

XX

XX 17-OCT-2000 (first entry)

XX

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 34526.

XX Hybridisation assay: genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway;  
 KW metabolic pathway; promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000: 2000EP-0301439.

XX PR 25-FEB-1999: 99US-0121825.

XX PR 05-MAR-1999: 99US-0123180.

XX PR 23-MAY-1999: 99US-0123548.

XX PR 23-MAY-1999: 99US-0125788.

XX PR 25-MAR-1999: 99US-0126264.

XX PR 01-APR-1999: 99US-0126785.

XX PR 06-APR-1999: 99US-0127462.

XX PR 16-APR-1999: 99US-0128234.

XX PR 15-APR-1999: 99US-0128714.

XX PR 21-APR-1999: 99US-0129845.

XX PR 23-APR-1999: 99US-0130077.

XX PR 23-APR-1999: 99US-0130449.

XX PR 25-APR-1999: 99US-0130510.

XX PR 25-APR-1999: 99US-0130891.

XX PR 30-APR-1999: 99US-0131449.

XX PR 30-APR-1999: 99US-0132049.

XX PR 01-MAY-1999: 99US-0132407.

XX PR 05-MAY-1999: 99US-0132484.

XX PR 05-MAY-1999: 99US-0132485.

XX PR 06-MAY-1999: 99US-0132486.

XX PR 07-MAY-1999: 99US-0132487.

XX PR 11-MAY-1999: 99US-0132863.

XX PR 14-MAY-1999: 99US-0134256.

XX PR 14-MAY-1999: 99US-0134218.

XX PR 14-MAY-1999: 99US-0134219.

XX PR 14-MAY-1999: 99US-0134221.

XX PR 18-MAY-1999: 99US-0134370.

XX PR 19-MAY-1999: 99US-0134768.

XX PR 20-MAY-1999: 99US-0134941.

XX PR 21-MAY-1999: 99US-0135124.

XX PR 21-MAY-1999: 99US-0135353.

XX PR 25-MAY-1999: 99US-0135629.

XX PR 26-MAY-1999: 99US-0136021.

XX PR 26-MAY-1999: 99US-0136792.

XX PR 01-JUN-1999: 99US-0137222.

XX PR 03-JUN-1999: 99US-0137528.

XX PR 04-JUN-1999: 99US-0137502.

XX PR 07-JUN-1999: 99US-0137724.

XX PR 08-JUN-1999: 99US-0138094.

XX PR 10-JUN-1999: 99US-0138540.

XX PR 10-JUN-1999: 99US-0138847.





```
Db 1181 G-CGCTGCCAGCAAGATAGCCCTGCTGGTCGGAGGGAICCCCTTCCTTGCTCG 1123
QY 320 atcattgttccactgtgccactgaagctgaccttcgcacctggagaggaggt 372
Db 1122 ATTTTGTCTTTCACATTAATCAACTGTGTCGGAGCTTTCACCTCGAGGGIGAT 1070

RESULT 15
AAC24234/c
ID AAC38234 standard: DNA: 485 BP.
XX
AC AAC38234;
XX
DT 17-OCT-2000 (first entry)
XX
DE Zea mays DNA fragment SEQ ID NO: 20250.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000: 2000EF-0301439.
XX
PR 25-FEB-1999: 99US-0121825.
PR 05-MAR-1999: 99US-0123180.
PR 09-MAR-1999: 99US-0123548.
PR 23-MAR-1999: 99US-0125788.
PR 25-MAR-1999: 99US-0126264.
PR 29-MAR-1999: 99US-0126785.
PR 01-APR-1999: 99US-0127452.
PR 06-APR-1999: 99US-0128234.
PR 08-APR-1999: 99US-0128714.
PR 16-APR-1999: 99US-0129845.
PR 19-APR-1999: 99US-0130077.
PR 21-APR-1999: 99US-0130449.
PR 23-APR-1999: 99US-0130510.
PR 28-APR-1999: 99US-0130891.
PR 30-APR-1999: 99US-0131449.
PR 30-APR-1999: 99US-0132048.
PR 04-MAY-1999: 99US-0132407.
PR 05-MAY-1999: 99US-0132484.
PR 06-MAY-1999: 99US-0132485.
PR 06-MAY-1999: 99US-0132486.
PR 07-MAY-1999: 99US-0132487.
PR 11-MAY-1999: 99US-0132863.
PR 14-MAY-1999: 99US-0134356.
PR 14-MAY-1999: 99US-0134218.
PR 14-MAY-1999: 99US-0134219.
PR 14-MAY-1999: 99US-0134221.
PR 18-MAY-1999: 99US-0134370.
PR 19-MAY-1999: 99US-0134768.
PR 20-MAY-1999: 99US-0134941.
PR 21-MAY-1999: 99US-0135124.
PR 21-MAY-1999: 99US-0135353.
PR 24-MAY-1999: 99US-0135629.
PR 25-MAY-1999: 99US-0136021.
PR 27-MAY-1999: 99US-0136392.
PR 28-MAY-1999: 99US-0136782.
PR 01-JUN-1999: 99US-0137222.
PR 03-JUN-1999: 99US-0137528.
PR 04-JUN-1999: 99US-0137502.
PR 07-JUN-1999: 99US-0137724.
PR 08-JUN-1999: 99US-0138094.
PR 10-JUN-1999: 99US-0138540.
PR 10-JUN-1999: 99US-0138847.
PR 14-JUN-1999: 99US-0139119.
PR 16-JUN-1999: 99US-0139452.
PR 16-JUN-1999: 99US-0139453.
PR 17-JUN-1999: 99US-0139492.
PR 18-JUN-1999: 99US-0139494.
PR 18-JUN-1999: 99US-0139454.
PR 18-JUN-1999: 99US-0139455.
PR 18-JUN-1999: 99US-0139456.
PR 18-JUN-1999: 99US-0139457.
PR 18-JUN-1999: 99US-0139458.
PR 18-JUN-1999: 99US-0139459.
PR 18-JUN-1999: 99US-0139460.
PR 18-JUN-1999: 99US-0139461.
PR 18-JUN-1999: 99US-0139462.
PR 18-JUN-1999: 99US-0139463.
PR 18-JUN-1999: 99US-0139750.
PR 18-JUN-1999: 99US-0139763.
PR 21-JUN-1999: 99US-0139817.
PR 22-JUN-1999: 99US-0139899.
PR 23-JUN-1999: 99US-0140353.
PR 23-JUN-1999: 99US-0140354.
PR 24-JUN-1999: 99US-0140695.
PR 28-JUN-1999: 99US-0140823.
PR 29-JUN-1999: 99US-0140991.
PR 30-JUN-1999: 99US-0141287.
PR 01-JUL-1999: 99US-0141842.
PR 01-JUL-1999: 99US-0142154.
PR 02-JUL-1999: 99US-0142055.
PR 06-JUL-1999: 99US-0142390.
PR 08-JUL-1999: 99US-0142803.
PR 09-JUL-1999: 99US-0142920.
PR 12-JUL-1999: 99US-0142977.
PR 13-JUL-1999: 99US-0143542.
PR 14-JUL-1999: 99US-0143624.
PR 15-JUL-1999: 99US-0144005.
PR 16-JUL-1999: 99US-0144085.
PR 16-JUL-1999: 99US-0144086.
PR 16-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144331.
PR 19-JUL-1999: 99US-0144332.
PR 19-JUL-1999: 99US-0144333.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144335.
PR 20-JUL-1999: 99US-0144352.
PR 20-JUL-1999: 99US-0144884.
PR 21-JUL-1999: 99US-0144814.
PR 21-JUL-1999: 99US-0145086.
PR 21-JUL-1999: 99US-0145088.
PR 22-JUL-1999: 99US-0145085.
PR 22-JUL-1999: 99US-0145087.
PR 22-JUL-1999: 99US-0145089.
PR 22-JUL-1999: 99US-0145192.
PR 23-JUL-1999: 99US-0145145.
PR 23-JUL-1999: 99US-0145218.
PR 23-JUL-1999: 99US-0145224.
PR 25-JUL-1999: 99US-0145276.
PR 27-JUL-1999: 99US-0145313.
PR 27-JUL-1999: 99US-0145318.
PR 27-JUL-1999: 99US-0145919.
PR 28-JUL-1999: 99US-0145951.
PR 02-AUG-1999: 99US-0146386.
PR 02-AUG-1999: 99US-0146388.
PR 03-AUG-1999: 99US-0146389.
PR 03-AUG-1999: 99US-0147038.
PR 04-AUG-1999: 99US-0147204.
PR 05-AUG-1999: 99US-0147302.
PR 05-AUG-1999: 99US-0147192.
PR 05-AUG-1999: 99US-0147260.
PR 06-AUG-1999: 99US-0147303.
PR 06-AUG-1999: 99US-0147416.
PR 09-AUG-1999: 99US-0147493.
PR 09-AUG-1999: 99US-0147935.
PR 10-AUG-1999: 99US-0148171.
PR 11-AUG-1999: 99US-0148319.
```



Result No.	Score	Query Match	Length	DB ID	Description
C 1	344.8	93.7	444	9	A1990990 wu35e01.x
C 2	343.2	93.3	388	9	AW584007 ia08e12.x
C 3	342.2	93.0	485	9	AW573307 EST385405
C 4	341.5	92.8	436	9	AA918966 o182e04.s
C 5	341	92.7	417	9	AA613829 nq26h02.s
C 5	340.6	92.6	383	9	AA983320 oq55hm4.s
C 7	340.2	92.4	399	10	BM030551 tgi5a06.y
C 8	335.8	91.2	426	9	AA918970 o182e08.s
C 9	329.8	89.5	385	10	BM055886 t8d4h11.y
C 10	329.8	89.6	715	10	BI835329 603085504
C 11	328.8	89.3	377	10	BG141336 ia33c04.y
C 12	325.4	88.4	395	9	AW582941 ia08b11.y
C 13	320.8	87.2	356	9	AA551894 nx13acs.s
C 14	319.8	86.9	718	10	BI836381 603083093
C 15	319.2	86.7	370	9	A1650251 w658h12.x
C 16	315.8	85.8	353	9	AA841948 ak61c03.s
C 17	314.8	85.5	318	9	AA845156 ak62f05.s



went through one round of normalization. Tissue samples provided by Dr. Brian Dieckgraefe (Washington University, dieck@im.wustl.edu); colonic mucosa represents a range of disease involvement from moderate to severe Crohn's disease; samples include both perforating (fistulas) and non-perforating samples. Library constructed by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 118 a 105 c 103 g 118 t  
ORIGIN

Query Match 93.7%; Score 344.8; DB 9; Length 444;  
Best Local Similarity 96.8%; Pred. No. 3.9e-95;  
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

QY 1 gaagagacgtgtaagtgcgggtcagattttcaactgaactctctggagcagagaacttcagcc 50  
|||||  
DB 384 GAAGAGACGTGTAAGTGGGGTGCAGCTTTCAACTGACCTCTGGAGCGCAGAACTTCAGCC 325  
|||||  
QY 61 atgaaggttaacagcgtctttcttcagtgcttgccttgccttgaagtccttctgttaac 120  
|||||  
DB 324 ATGAAGGTAAACAGCGATCTTCTTCAGTGGCTTGGCCCTTGAGTCTATCIGGTAAC 255  
|||||  
QY 121 actggagctgaactccctggagagagggccaaatgttacaatgaacttaagtgaaccc 180  
|||||  
DB 254 ACTGGAGCTGACTCCCTGGAGAGAGAGGCCAAATGTACAACTTAATGCAATGCACC 205  
|||||  
QY 181 agatatatagacctgctgtggagactgatgaataacttaccatgaatgaatgcgttta 240  
|||||  
DB 204 AAGATATAGACCTGTCTGTGGGACGTGAGGAATACITATCCCAATCAATTCGGTGTA 145  
|||||  
QY 241 tgtttgaaggtcggaacacccggagcttctatctcattcaaaatctgggcttactga 300  
|||||  
DB 144 TGTITGAAAATCGGAACGCGAGACTTCTATCCICATTCNAAAATCTGGGCTTGCIGA 85  
|||||  
QY 301 gaaccaaggtttgaatccatccatcaggtcacccagagcc-----tattgtttaa 350  
|||||  
DB 84 GAACCAAGGTTTGAATCCCATCAGGTACCCGAGAGGCTGACTGGCCCTTATGTGTA 25  
|||||  
QY 351 taatgtatctgaatc 358  
|||||  
DB 24 TAAATGATCTGAATATC 7  
|||||

RESULT 2  
AW584007/c  
LOCUS  
DEFINITION  
gb:M11949 PANCREATIC SECRETORY TRYPsin INHIBITOR PRECURSOR (HUMAN);  
gb:X06342 Mouse mRNA for prostatic secretory glycoprotein (MOUSE);  
mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
human.  
Homo sapiens

REFERENCE  
AUTHORS  
Melton, D., Brown, J., Kenty, G., Permut, A., Lee, C., Kaestner, K.,  
Lemishka, I., Scaer, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blais, A.,  
Schmitt, A., Theising, B., Ritter, E., Ponko, I., Bennett, J., Cardenas,  
M., Gibbons, M., McCann, R., Cole, R., Isagareishvili, R., Williams, I.,  
Jackson, Y. and Boxers, Y.

TITLE  
JOURNAL  
COMMENT  
Unpublished (2000)  
Other ESTs: ia08el2.yl  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept. of Molecular and Cellular Biology, 7 Divinity Ave. Cambridge,  
MA 02138

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 388)  
MELTON, D., BROWN, J., KENTY, G., PERMUT, A., LEE, C., KAESTNER, K.,  
LEMISHKA, I., SCAER, M., BRESTELLI, J., GRADWOHL, G., CLIFTON, S.,  
HILLIER, L., MARRA, M., PAPE, D., WYLIE, T., MARTIN, J., BLAIS, A.,  
SCHMITT, A., THEISING, B., RITTER, E., PONKO, I., BENNETT, J., CARDENAS,  
M., GIBBONS, M., MCCANN, R., COLE, R., ISAGAREISHVILI, R., WILLIAMS, I.,  
JACKSON, Y. and BOXERS, Y.

Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmel@conbioh.harvard.edu  
Library was constructed by Dr. Douglas Melton DNA sequencing by:  
Washington University Genome Sequencing Center For information on  
obtaining a clone please contact: Juliana Brown  
(brownjfas.harvard.edu)  
Seq primer: 40up from Gibco.

FEATURES  
Location/Qualifiers  
Source  
1..388  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="Human Pancreatic Islets"  
/tissue\_type="Islets of Langerhans"  
/dev\_stage="Adult"  
/lab\_host="DH10B"  
/note="Organ: Pancreas; Vector: pSPORT1; Site: 1; Not 1;  
Site: 2; Sal 1; Library constructed using SuperScript  
Plasmid Library kit (Life technologies). cDNA made by  
oligo-dt priming. Size-selected by column fractionation;  
average insert size 1.08 kb. Primary library,  
unamplified."

BASE COUNT 105 a 92 c 84 g 106 t  
ORIGIN

Query Match 93.3%; Score 343.2; DB 9; Length 388;  
Best Local Similarity 96.6%; Pred. No. 3.9e-95;  
Matches 365; Conservative 0; Mismatches 3; Indels 10; Gaps 1;

QY 1 gaagagacgtgtaagtgcgggtcagattttcaactgaactctctggagcagagaacttcagcc 50  
|||||  
DB 391 GAAGAGACGTGTAAGTGGGGTGCAGCTTTCAACTGACCTCTGGAGCGCAGAACTTCAGCC 322  
|||||  
QY 61 atgaaggttaacagcgtctttcttcagtgcttgccttgccttgaagtccttctgttaac 120  
|||||  
DB 321 ATGAAGGTAAACAGCGATCTTCTTCAGTGGCTTGGCCCTTTCAGTCTATCIGTAAAC 262  
|||||  
QY 121 actgaagctgaactccctggagagagggccaaatgttacaatgaacttaagtgaaccc 180  
|||||  
DB 251 ACTGGAGCTGACTCTCTGGAGAGAGGCCAAATGTACAACTTAATGGAATTCAGCC 202  
|||||  
QY 181 agatatatagacctgctgtggagactgatgaataacttaccatgaatgaatgcgttta 240  
|||||  
DB 201 AAGATATAGACCTCTCTGGAGACTGATGCAATACITATCCCAATGAATGCGTGTA 142  
|||||  
QY 241 tgtttgaaggtcggaacacccagagcttctatctcattcaaaatctgggcttactga 300  
|||||  
DB 141 TGTITGAAAATCGGAACGCGCAGACTTCTATCCTCATTCAAAATATATGGCCTTGCCTGA 82  
|||||  
QY 301 gaaccaaggtttgaaatcccatcaggtcacccagagcc-----tattgtttaa 350  
|||||  
DB 81 GAACCAAGGTTTGAATCCCATCAGGTACCCGAGAGGCTGACTGGCCCTTATGTGTA 22  
|||||

QY 351 taatgtatctgaatc 358  
|||||  
DB 21 TAAATGATCTGAATATC 4  
|||||

RESULT 3  
AW973307/c  
LOCUS  
DEFINITION  
EST1395405 MAGE resequences, MAGM Homo sapiens cDNA, mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
human.  
Homo sapiens

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 485)  
HEDGE, P., OI, R., ABERNATHY, K., DHARAP, S., GASPARD, R., GAY, C., HOLT,  
J. E., SAED, A. I., SHAROV, V., LEE, N. H., YEATMAN, T. J. and



```

VERSION      AA612829.1  GI:2453857
KEYWORDS     EST.
SOURCE       Homo sapiens
ORGANISM     Homo sapiens
REFERENCE    1 (bases 1 to 417)
AUTHORS     Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE       Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT     National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
            Tumor Gene Index
            Unpublished (1997)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs@mail.nih.gov
            Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D.,
            Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www.bio.llnl.gov/bbrp/image/image.html
            Insert Length: 758 Std Error: 0.00
            Seq primer: -40ml3 fwd. ET from Amersham
            High quality sequence stop: 395.
            Location/Qualifiers
            source
            1..417
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:1145057"
            /tissue_type="colon tumor RER"
            /lab_host="DH10B"
            /note="Organ: colon; Vector: pT73D-PAC (Pharmacia) with a
            modified polylinker; 1st strand cDNA was prepared from
            RER+ colon tumor, and was then primed with a Not I -
            oligo(dT) primer. Double-stranded cDNA was ligated to Eco
            RI adaptors (Pharmacia), digested with Not I and cloned
            into the Not I and Eco RI sites of the modified pT73
            vector. Library is normalized. Library was constructed by
            Bento Soares and M. Fatima Bonaldo (N-Soares4)."
            BASE COUNT  115 a  96 c  97 g  109 t
            ORIGIN
            Query Match      92.7%; Score 341; DB 9; Length 417;
            Best Local Similarity 95.3%; Pred. No. 1.9e-94;
            Matches 363; Conservative 0; Mismatches 5; Indels 9; Gaps 1;
            QY 1 gaagagcgtggaagtcgagtcagcttctcaactcaactcctgagacacgaacttcagcc 60
            Db 377 GAAGAGCGTGGTAAGTGGCGGTCAGTTCACCTGACCTCTGGACGACGACCTTCAGCC 318
            QY 61 atgaagataacaggcatcttctctcagtcgcttgccctgttgatctatctgtatac 120
            Db 317 ATGAAGGTAACAGGCATCTTCTCAGTCGCTTGGCCCTGTGAGTCTATCTGTGTAAC 258
            QY 121 actgagactgactccctgggaagagagacaaatgtttacaaatgaacttaagtgcacc 190
            Db 257 ACTGGAGCTGACTCCCTGGGAAGAGAGGCGCAATGTGTACATGAACCTTAATGGAATGACC 198
            QY 181 aagatatatacctctgtctggagactgaggaataacttaccatgaatgcgtgatta 240
            Db 197 AAGATATATGACCTGTCTGTGGGACTGATCGAATACTATCCCAATGAATGCGTGTA 139
            QY 241 tgttttaagtcggaagcgcagactctctatcctcattcaaaaatctggcccttcgga 300
            Db 137 TGTITTTGAAAATCGGAACGCGACACTCTATCCTCATTCAAAATCTGGGCGCTTGCTGA 78
            QY 301 gaaccaaggtttgaaatcccatcagatcacccgga-----gccttatgtgaat 351
            Db 77 GAACCAAGGTTTTGAATCCCATCGGTCACTGCGAGCCCTGACTGCGCTTATTTGTTGAA 18
            QY 352 aaatgtatctgaatc 368

```

```

Db 17 AAAATGTAICTGAATC 1
|||||
AA983320 383 bp mRNA linear EST 23-JUL-1999
Cq56h04.s1 NCI-CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1590391 3'
similar to gb:M11949 PANCREATIC SECRETORY TRYPSIN INHIBITOR
PRECURSOR (HUMAN);, mRNA sequence.
AA983320
AA983320.1 GI:3151845
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 383)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 448 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 252.
            Location/Qualifiers
            source
            1..383
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:1590391"
            /tissue_type="NCI-CGAP_Kid5"
            /lab_host="DH10B"
            /note="Organ: kidney; Vector: pT73D-PAC (Pharmacia) with
            a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
            strand cDNA was primed with a Not I - oligo(dT) primer [5].
            AACTGGAAGAAATTCGCGCGCAATATTTTITTTTTTTT 3'].
            double-stranded cDNA was ligated to Eco RI adaptors
            (Pharmacia), digested with Not I and cloned into the Not I
            and Eco RI sites of the modified pT73 vector. Library
            went through one round of normalization. Library
            constructed by Bento Soares and M. Fatima Bonaldo.
            BASE COUNT  103 a  88 c  85 g  107 t
            ORIGIN
            Query Match      92.6%; Score 340.6; DB 9; Length 383;
            Best Local Similarity 95.5%; Pred. No. 2.4e-94;
            Matches 362; Conservative 0; Mismatches 4; Indels 9; Gaps 1;
            QY 3 agagacgtggaagtcgagtcagcttctcaactgacctctggacgcagaaacttcagccat 62
            Db 383 AGAGACGTGTAAGTGGGTGCAGTTTCAACTGACCTCTGGACGACGAACTTCAGCCAT 324
            QY 63 aaagataacaggcatcttctctcagtcgcttgccctgttgatctatctgtatacac 122
            Db 223 GAAGGTAACAGGCAATCTTCTCTCAGTGGCTTGGCCCTGTGAGTCTATCTGTGTAAC 264
            QY 123 tggagctgactccctgggaagagagccaaatgtttacaaatgaacttaagtgcaccaa 182
            Db 263 IGAGAGTGAATCCCTGGGAAGAGAGGCGCAATGTGTACATGAACCTTAATGATGACCAA 204
            QY 183 tatatatgacctgtctgtgtgggactgatggaataacttaccatgaatgcgtgtatg 242
            -----

```

```

|||||
Db 203 GATATATGACCCCTGTCTGTGGGACTGATGAATACTATATCCCAATGAATCGGTGTAIG 144
QY 243 tttaaaggctggaacgacgactctctcctcattcaaaaatctgggcttgcctgaaga 302
|||||
Db 143 TTTTGAATAATCGGAACGCCAGACACTTCTATCTCTATTCATTCATCAAAATCTGGGCCCTGCTGAGA 84
QY 303 accaaggttttgaataatccctacaggtacacccaga-----ggcctattttgaataa 353
|||||
Db 83 ACCAAGGTTTGTGAATCCCAATGAGGTACCGGAGCGCTGACTGGCCCTTATTTGTGATAA 24
|||||
QY 354 atgtatctgaatctc 368
Db 23 ATGTATCTGAATATC 9
|||||

RESULT 7
BM090551 399 bp mRNA linear EST 20-NOV-2001
LOCUS
DEFINITION
ig15406.y1 Human Fetal Pancreas 1A Homo sapiens cDNA 5' similar to
SW:IFS1.HUMAN F00995 PANCREATIC SECRETORY TRYPsin INHIBITOR
PRECURSOP : mRNA sequence.
BM090551
ACCESSION
VERSION
BM090551.1 GI:17019517
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 399)
MELTON,D., BROWN,J., KENTY,G., PERMUTT,A., LEE,C., KAESTNER,K.,
LEMISHKA,I., SCEARCE,M., BRESTELLI,J., GRADWOHL,G., CLIFTON,S.,
HILLIER,L., MARIA,M., PAPE,D., WYLIE,I., MARTIN,J., BLISTAIN,A.,
SCHMITT,A., THEISINGER,B., RITTER,E., KONKO,I., BENNETT,J., CARDENAS
,M., GIBBONS,M., MCCANN,R., COLE,R., TSAGAREISHVILI,R., WILLIAMS,T.,
JACKSON,Y. and BOWERS,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center for information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@imgate.wustl.edu)
Putative full length read
vector to vector length is.
FEATURES
Location/Qualifiers
1..399
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human Fetal Pancreas 1A"
/tissue_type="Fetal Pancreas (4 Pooled Donors, 18 - 20
weeks, Stratagene #738023)".
/db_xref="Fetal Pancreas"
/note="vector: pBluescript SK(-); Site1: NotI; Site2:
XhoI; cDNA made by oligo-dT priming. Size selected by
CFPCMA SPIN-1000 (Clontech). Average insert size -1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110 E-mail: hinoue@imgate.wustl.edu Tel:
314-362-1916, Fax:314-747-2692."
BASE COUNT
116 a 86 c 92 g 105 t
ORIGIN

```

```

Query: Match 92.4%; Score 340.2; DB 10; Length 399;
Best Local Similarity 96.5%; Pred. No. 3.3e-94;
Matches 362; Conservative 0; Mismatches 3; Indels 10; Gaps 1;

QY 4 gacacgtgtaagtcggtgcagatttttcaactgacctctggacgcagaaacttcagccatg 63
|||||
Db 11 GAGACGTGGTAAAGTGGGTGCAGTTTCAACTGACCTCTGGACGAGAACTTCAGCCATG 70
QY 64 aaggttaacagggatctttcttcagtgctggcccttgaagtcctatctatgtaaacatc 123
|||||
Db 71 AAGGTAAACAGGCACTTCTCTCAGTGTCTGGCCCTGTGAGTCTATCTGGTAACACT 130
QY 124 gaaactgactcctcggagaagaagccaaatgttacaatgaacttaaatgaatgacccaag 183
|||||
Db 133 GGAGCTGACTCCCTGGGAAGAGAGGCCAATGTTACAACTTAATGGATGCACCAAG 190
QY 134 atatagacccctgtctgtggactgagtgaaatacttatcccaatgaatgcgtgttatgt 243
|||||
Db 151 ATATATGACCCCTGTCTGTGGGACTGATGGAATACTATCCCAATGAATCGGTGTAIG 250
QY 243 tttaaaggctggaacgacgactctctcctcattcaaaaatctgggcttgcctgaaga 303
|||||
Db 251 TTTGAAAATCGGAACGCCAGACTTCTATCTCTCAATCAAAATCTGGGCCCTGCTGAGA 310
QY 304 ccaaggttttgaataatccctacaggtcacccgagcc-----tattgttgaataa 353
|||||
Db 311 CCAAGGTTTTGAAATCCCATCAGTCACGTCGAGGCGCTGACTGGCCCTTATTTGTGATAA 370
QY 354 atgtatctgaatctc 368
Db 371 ATGTATCTGAATATC 385
|||||

RESULT 8
AA918970 426 bp mRNA linear EST 10-JUN-1999
LOCUS
DEFINITION
ol82908.s1 NCI_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1536158 3'
similar to gb:M11949 PANCREATIC SECRETORY TRYPsin INHIBITOR
PRECURSOP (HUMAN):: mRNA sequence.
AA918970
ACCESSION
VERSION
AA918970.1 GI:3058860
KEYWORDS
ESI.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 426)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert length: 623 Std Error: 0.00
Seq primer: -40m13 fwd. Et from Amersham
High quality sequence stop: 409.
FEATURES
Location/Qualifiers
1..426
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:1536158"
/clone_lib="NCI_CGAP_Kid5"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with

```

a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-AACTGAGAAATCGCGCCGCAATATTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 114 a 97 c 100 g 115 t

Query Match 91.2%; Score 335.8; DB 9; Length 426;

Best Local Similarity 95.3%; Pred. No. 7.7e-93;

Matches 361; Conservative 0; Mismatches 7; Indels 11; Gaps 1;

QY 1 gaagagagctggttaagtgcagtcagtttcaactgaacctctggagcagaaacttcagcc 60  
 Db 386 GAAGAGAGCTGGTAAGTGGCGGTCAGTTCACACTGACCTCTGGAGCGCAACTTCAGCC 327  
 QY 61 ataaagataacagagcatctttctcaatgccttgagcctgttgagttatctatgaac 120  
 Db 326 ATGAAGTAAACAGCATCTTCTTCTCAGTGGCTTGGCCCTGTGAGTCTATCTGGTAAC 257  
 QY 121 actggagctgaactccctggaggaagagggccaaatgttacaataaacttaatgatcac 180  
 Db 266 ACTGGAGCTGACTCCCTGGGAGAGAGGCCAAATGTTACAACTTAATGGATGCACC 207  
 QY 181 agatatataacacctctctgtggagctgatgaatactatccccaatgaatgcgttta 240  
 Db 206 AAGAATATGACCTCTGTCGTGGAGCTGATGGAATACTTATCCCAATGAATGCGTGTA 147  
 QY 241 tgtttgaagctggaagcgcagactctctactctcattcaaaaactctggccttgcga 300  
 Db 146 TGTTTTAAAAICGGAACGCCAGACTTCTATCTTCATTCAAAAATCTGGGCTTGCCTGA 87  
 QY 301 gaaccaaggtttgaataaccatcaggtcaccc-----gcgagggcctattgttga 349  
 Db 86 GAACCAAGGTTTTGAATCCCATCAGGTCAAGTCACCCGCGGAGCCCTGACTGGCTTATTTGA 27  
 QY 350 ataatgtatctgaatc 368  
 Db 26 ATAAATGATCAGAAATC 8

# RESULT 9

## BM055886

### LOCUS

DEFINITION id84bl1.y1 Melton Normalized Mixed Mouse Pancreas 1 N1-MMS1 Mus musculus cDNA 5' similar to SW:IPSI\_HUMAN P00995 PANCREATIC SECRETORY TRYPsin INHIBITOR PRECURSOR 1; mRNA sequence.

## ACCESSION

### VERSION

### KEYWORDS

### SOURCE

### ORGANISM

Mus musculus house mouse.  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

### AUTHORS

Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K., Lemshka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A., Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, R., Williams, I., Jackson, Y., and Bowers, Y.  
 Endocrine Pancreas Consortium  
 Unpublished (2000)

## TITLE

### JOURNAL

### COMMENT

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept. of Molecular and Cellular Biology, 7 Divinity Ave. Cambridge, MA 02138

Tel: 617-495-1812  
 Fax: 617-495-8557

Email: dmelton@bioche.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Juliana Brown (brownj@fas.harvard.edu).

## FEATURES

### SOURCE

1..385

/organism="Mus musculus"

/strain="ICP"

/db\_xref="taxon:10090"

/clone\_lib="Melton Normalized Mixed Mouse Pancreas 1 N1-MMS1"

/sex="Both for embryonic & newborn, male for adult and adult islet"

/dev\_stage="Embryonic day 10.5, E12.5, E16.5, newborn, adult, mixed"

/lab\_host="DH10B"

/note="vector: pSPOT1; Site\_1: Not I; Site\_2: Sal I; Five libraries representing E10.5/12.5 pancreatic bud, E16.5 pancreas, newborn pancreas, adult pancreas, and adult islets of Langerhans were separately constructed using SuperScript Plasmid Library kit (Life Technologies). cDNA was made by oligo-dT priming and size-selected by column fractionation. Libraries were amplified once on solid support and plasmid DNA from each library was prepared and mixed in equal amounts. The mixed library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded mixed library plasmid DNA was mixed with 5 micrograms PCR product representing mixed library inserts and hybridized to an EcoT of 6. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

BASE COUNT 113 a 81 c 84 g 101 t

## ORIGIN

Query Match 89.6%; Score 329.8; DB 10; Length 385;

Best Local Similarity 96.7%; Pred. No. 5.2e-91;

Matches 351; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

QY 16 gtcgggtcagatttcaactgaacctctggagcagaaacttcagccatgaagtaacagcc 75  
 Db 1 GTGCGGTCAGTTCACACIGACCTTCGACGACGACCTTCAGCCATGAAGTAACAGGC 60  
 QY 76 atctttcttcagtcgcttgagcctgttgagctctatctgtaaacactgagctgactcc 135  
 Db 81 ATCTTCTTCTCAGTGGCTTGGCCCTGTGTAGTCTATCTGTAACTGAGCTGACTCC 120  
 QY 136 ctgggaagagagagccaaatgttacaataaacttaatgagtcacccaagatatatgacct 195  
 Db 121 CTGGGAGAGAGAGGCCAAATGTTACAACTTAATGATGCACCAAGATATATGACCT 180  
 QY 196 atctgtggagctgatgaataacttaccgaatgaatgcgtgttatgtttgaagctgg 255  
 Db 191 GTCTGGGACTGATGGAATTAATCTCCCAATGAATCGCTGTATGTTTGAATTCGG 240  
 QY 256 aaacgacagactctatctctcattcaaaaactctggccttgctgagaccaaagtttga 315  
 Db 241 AAACGCCAGACTTCTATCTTCATTCANAAATCTGGGCTTGTCTGAGAACCAAGTTTGA 300  
 QY 316 atcccatcaggtcacccgcaggccc-----tattgttgataaatgtatctgaat 365  
 Db 301 AATCCCATCAGGTCACCGCGAGGCTGACTGGCCCTTATTGTTGAATAAATGATCTGAAT 360

366 atc 368

367 att

368 atc 363

## RESULT 1

```

B1835329      715 bp      mRNA      linear      EST 04-OCT-2001
LOCUS          50308504F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5224728 5'.
DEFINITION     mRNA sequence.
ACCESSION      B1835329
VERSION        B1835329.1 GI:15946879
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 715)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: coapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM1565 row: g column: 01
High quality sequence stop: 449.
Location/Qualifiers
1..715
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5224728"
/clone_lib="NIH_MGC_120"
/lab_host="DH10B"
/note="Organ: pooled pancreas and spleen; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (In: Nitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC Library."
BASE COUNT    251 a 166 c 190 g 108 t
ORIGIN

Query Match      89.5%; Score 329.8; DB 10; Length 715;
Best Local Similarity 96.7%; Pred. No. 6.7e-91;
Matches 351; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

Q: 16 gtgcgtgcagttttcaactgacctctggacgcagaacttcagccatgaaggttaacaggc 75
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D: 1 GTGCGGTGCAGTTTTCACACIGACCTCTGGACGCCAGAACTTCAGGCATGAAGGTAAACAGGC 60

Q: 76 atctttcttcagtgcttgccctgttgagttctatctgtatacactgagctgactcc 135
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D: 61 ATCTTTCTCTCAGTGCCTTGCGCCGTGTGTGAGTCTATCTGTTAACTTGGACCTGACCTCC 120

Q: 136 ctgggaagagagcccaatgtttacataaacttaataatgagatgaccacaaatataatgacctt 195
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D: 121 CTGGGAGAGAGCCCAATGTTTACAACTTAATGATGATGATGATGATGATGATGATGATGATGAT 180

Q: 196 gctgtgggaactgataagaaatacttatcccaatgaatgcgtgttatgttttgaagggtcgg 255
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D: 181 GTCTGTGGGACTGATGGAATACTTATCCCAATGAATGCGTGTATGTTTGTGAAAATCGG 240

Q: 256 aaacgcgaactctatctcctcattcaaaaactctggcccttgcctgaagaaccaaggtttga 315
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D: 241 AAACGCCAGACTTCTATCTCTCAATCAAAATCTGCGCTTCTGAGAACCAAGGTTTTGA 300

Q: 316 aatcccatcaggtcacccgcgagcc-----tattgttgataaataatgtatctgaat 365
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D: 301 AATCCCATCAGGTCAACCGCGAGGCTGACTGCGCTTATGTGTGATAAATGATGATGATGAT 350

```

Q: 346 atc 368  
Db 361 ATC 363

# RESULT 11

BG141336

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..377

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="Melton Human Islets H121"

/sex="Both"

/tissue\_type="Islets of Langerhans"

/dev\_stage="Adult"

/lab\_host="TOP10"

/note="Organ: Pancreas; Vector: pZEO-2; Site\_1: Not I; Site\_2: Xho I; Library constructed using SuperScript Plasmid Library kit (Life Technologies). cDNA made by oligo-dT priming. Xho I site destroyed during cloning. Size-selected by column fractionation; average insert size 1.59 kb. Primary library, unamplified."

BASE COUNT 112 a 81 c 83 g 101 t

ORIGIN

Query Match 89.3%; Score 328.8; DB 10; Length 377;

Best Local Similarity 96.7%; Pred. No. 1.1e-90;

Matches 350; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

Q: 17 tgcggtagcttttcaactgacctctggacgcagaacttcagccatgaaggttaacaggc 76

Db 1 TCGGTGCAGTTTTCACACTGACCTCTGGACGCCAGAACTTCAGGCATGAAGGTAAACAGGC 60

Q: 77 tcttctcttcagtgccctgttgagttctatctgttaacactggaggtgactccc 135

Db 71 TCTTCTCTCAGTGCCTTGCCCTGTGTGAGTCTATCTGTTAACTTGGACCTGACCTCC 120

Q: 137 tgggaagagagcccaatgtttacataaacttaataatgagatgaccacaaatataatgacctg 195

```

|||||
Db 121 TGGAGAGAGGCGCAATGTTACAAAGCTTAATGATGACCAAGATAATACCCCTG 180
Oy 197 tctgtggactgatggaaataactatcccaatgaatcgtgttatgtgttttgaaggtcaga 256
|||||
Db 181 TCTGTGGAGCTGATGGAATACATATCCCAATGNAATGCGTGATGTGTTTGAAATCGGA 240
Oy 257 aacccagacttctatcctcattcaaaaattcgggcttctctgagaaaccaaggtttttaa 315
|||||
Db 241 AACCCAGAGCTTATCCATCAITCAAAAATCTGGGCTTCTGAGAACCAAGGTTTGAA 300
Oy 317 atcccatcaggtcaccgcggggc-----tattgttgtaataatctatctgaata 365
|||||
Db 301 ATCCCATCAGTCACCGGAGGCTGCTGCGCTTATGTTGTAATAATGATCTGAATA 360
Oy 367 tc 358
Db 361 TC 362

RESULT 12
AW582941
LOCUS
DEFINITION
  1a08b11.71 Human Pancreatic Islets Homo sapiens cDNA, 5' similar to
  qb:M11949 PANC-PEATIC SECRETORY TRYPsin INHIBITOR PRECURSOR (HUMAN);
  qb:X05342 Mouse mRNA for prostatic secretory glycoprotein (MCUSE);.
  mRNA sequence.
ACCESSION
  AW582941
VERSION
  AW582941.1 GI:7259985
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 395)
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, E., Panko, I., Bennett, J., Cardenas
, M., Gibbons, M., McGann, R., Cole, R., Tsagaris, R., Williams, J.,
Jackson, Y., and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University; Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brownjfas.harvard.edu)
Seq primer: -40RP from Gibco.
1. 395
Location/Qualifiers
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone_lib="Human Pancreatic Islets"
  /tissue_type="Islets of Langerhans"
  /dev_stage="Adult"
  /lab_host="DH10B"
  /note="organ: Pancreas; Vector: pSPORT1; Site: 1; Not 1;
  Site 2: Sal 1; Library constructed using SuperScript
  Plasmid Library kit (Life technologies). cDNA made by
  oligo-dt priming. Size-selected by column fractionation:
  average insert size 1.02 kb. Primary library,
  unamplified."
105 a 87 c 96 g 107 t

BASE COUNT
ORIGIN

```

```

Query Match      88.4%; Score 325.4; DB 9; Length 395;
Best local Similarity 95.5%; Pred. NO. 1.2e-89;
Matches 360; Conservative 0; Mismatches 6; Indels 11; Gaps 2;

Oy 1 gaagagacgtggaagtgcggtgcagttttcaactaacctcttgacgcagaaacttcagcc 60
Db 19 GAAGAGACGTGGTGAAGTGGCGTGCAGTTCCTCAACTGACCTCTGAGCGCAGAACTTCAGCC 78
Oy 61 atcaaggttaacagagctctttcttctcagtcagtccttgacccctgttgagttctat-ctggtaa 119
Db 79 ATGAAGGTAACAGAGCACTTCCTTCAGTGCCTTGCCCTGTGAGCTATGAGGGGAA 138
Oy 120 cactgagctgactccctcctgggaagagcgaatgttacaatgaacttaagtgaatcac 179
Db 132 CACTGAGCTGACTCCCTGGGAGAGAGGCAAAATGTTACATGAACCTTAATGGATGCAC 198
Oy 190 caagatatatgacctctctgtgggactgatgaaatactatcccaatgaatgcctgtt 239
Db 193 CAAGAATATGACCTGTCTGTGGGACTGATGGAATACTTATCCCAATGAATGCGGT 258
Oy 240 atgttttgaaggtcgaagaaacccaaactctctcattcattcaaaaatctgggcttgc 299
Db 259 ATGTTTGAATAATCGAAACGCCAGACTTCTATCCTCATITCAAAAATCTGGGCTTGTG 318
Oy 300 aaaaacaaagttttgaaatcccatcaggtccacgcgagggc-----tattattga 349
Db 312 AGAACCAAGGTTTIGAAATCCCATCATGTCACCGCGAGGCGCTGACGTGGCTTATTGTGA 378
Oy 350 ataaatgtatctgaata 366
Db 353 ATAAATGATCTGAATA 395

RESULT 13
AW551894/C
LOCUS
DEFINITION
  Similar to gb:M11949 PANC-REATIC SECRETORY TRYPsin INHIBITOR
  PRECURSOR (HUMAN);. mRNA sequence.
ACCESSION
  AW551894
VERSION
  AW551894.1 GI:2322146
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 356)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc., David B. Krizman,
Ph.D.
cDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 457 Std Error: 0.00
Seq primer: -40m13 fwd. Fr from Amersham.
1. 356
Location/Qualifiers
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone_lib="IMAGE:1013360"
  /tissue_type="tumor"
  /lab_host="SOLR (kanamycin resistant)"

```





```
/issue_type="colonic mucosa from 3 patients with Crohn's
disease"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pTZ19-Pac (Pharmacia) with a
modified polylinker; Site: 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'-
IGTACCAATCTGAAGGAGGCGGCGGCTTTTTTTTTTTTTTTT 3']
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pTZ19 vector. Library
went through one round of normalization. Tissue samples
provided by Dr. Brian Dieckgraefe (Washington University,
St. Louis, MO; bdieck@wustl.edu); colonic mucosa represents a range of
disease involvement from moderate to severe Crohn's
disease; samples include both perforating (fistulas) and
non-perforating samples. Library constructed by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT      105 a      98 c      78 g      98 t
ORIGIN
```

```
Query Match      86.7%      Score 319.2: DB 9: Length 370:
Best Local Similarity 93.8%      Pred. No. 9.5e-88:
Matches 347: Conservative 0: Mismatches 13: Indels 10: Gaps 1:

Oy 3 agagacgtgataagtcggtgcagattttcaactaacctctggagcagcaaaccttcagccat 52
Db 370 AGAGACGTGTAGTGGGGTGGCGGTTCAGCTTCAACTGTCTCTGTACGAGAACCTTCACCCAT 311

Oy 63 gaaggtaacaggcattcttcttcagtcgcttggccctgttgaagtctatctggtaacac 122
Db 310 GAAGGTAACAGGCACTTCTTATCAGTGCCTTGGCCCTGTTCAGCTATCTGGTAACAC 251

Oy 123 tgaagctgaactccctgagaaagagaccacaattttacaatgaacttaattgagatgacccaa 182
Db 250 TGAAGCTGACTCCCTGGGAGAGAGAGGCCAAATGTTCATGAACCTTAATGGATGCACCAA 191

Oy 183 gatatagaccctgtctgtgggactgatgaaataacttaccatgaatgagtggtttatg 242
Db 190 GATATATGACCCGTCTGTGGGAATGATGAATAACTTATCCCAATGAATGCGTGTATG 131

Oy 243 ttltgaaggctcgaagacgcagacacttctatctcattcaaaaaatctgggaccttgctgaga 302
Db 130 TTTTGAATAATCGAAACGGCAGCGCTTCTATCTCTCATTCAAAAAATTTGGGCTTGTGAGA 71

Oy 303 accaaggcttttgaatcccatcagggtcaccgcagagcc-----tatgtttgaata 352
Db 70 ACCAAGGTTTGAATAATCCCATCAGGTCTCCGCGAGGCGCTTATGTTGAATA 11

Oy 353 aatgtatctg 352
Db 10 AATGTATCTG 1
```

Search completed: July 29, 2002, 17:20:25  
Job time: 7745 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 19:27:19 ; Search time 394.41 Seconds  
(without alignments)  
1967.976 Million cell updates/sec

Title: US-09-880-107-3847

Perfect score: 368  
Sequence: 1 gaagagacgtgtaagtgcg.....aataaatgtatcgaatc 368

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1333638 seqs, 1054605264 residues

Total number of hits satisfying chosen parameters: 2667276

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Pending\_Patents\_NA\_Nev.\*  
1: /cqn2\_6/ptodata/2/pna/PCT\_NEW\_COMB.seq.\*  
2: /cqn2\_6/ptodata/2/pna/US06\_NEW\_COMB.seq.\*  
3: /cqn2\_6/ptodata/2/pna/US07\_NEW\_COMB.seq.\*  
4: /cqn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq.\*  
5: /cqn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq.\*  
6: /cqn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq.\*  
7: /cqn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq.\*  
8: /cqn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	344.8	93.7	603	7	US-10-162-749-16
2	344.8	93.7	607	7	US-10-162-749-57
3	210.8	57.3	214	8	US-60-382-445-457
4	88.8	24.1	180	5	US-09-454-226A-557
5	45	12.2	421	8	US-60-365-384-494
6	45	12.2	652	8	US-60-365-384-156
7	42.2	11.5	323	1	PCT-US01-43704-206
8	42.2	11.5	323	1	PCT-US01-43704-2330
9	42.2	11.5	398	5	US-09-671-483-2
10	42.2	11.5	433	7	US-10-106-698-1897
11	40.6	11.0	3528	1	PCT-US02-10824-19
12	40.6	11.0	3532	5	US-09-926-820-7
13	40.4	11.0	156	5	US-09-926-820-12
14	39.2	10.7	562	5	US-09-629-469A-8198
15	39.2	10.7	1022	5	US-09-629-469A-11819
16	38.2	10.4	180	5	US-09-671-483-1
17	38	10.3	338	1	PCT-US02-10421-1433
18	38	10.3	338	7	US-10-112-699-1433
19	36.6	9.9	365	1	PCT-US02-10421-1457
20	36.6	9.9	365	7	US-10-112-699-1457
21	36.6	9.9	394	1	PCT-US02-10421-1212
22	36.6	9.9	394	7	US-10-112-699-1212
23	36.6	9.9	395	1	PCT-US02-10421-2209
24	36.6	9.9	395	7	US-10-112-699-2209
25	36.6	9.9	426	1	PCT-US02-10421-1089

26	36.6	9.9	426	7	US-10-112-699-1089	Sequence 1089, Ap
27	36.6	9.9	437	1	PCT-US02-10421-2071	Sequence 2071, Ap
28	36.6	9.9	437	7	US-10-112-699-2071	Sequence 2071, Ap
29	36.6	9.9	439	1	PCT-US02-10421-1527	Sequence 1527, Ap
30	36.6	9.9	439	1	PCT-US02-10421-2125	Sequence 2125, Ap
31	36.6	9.9	439	7	US-10-112-699-1527	Sequence 1527, Ap
32	36.6	9.9	439	7	US-10-112-699-2125	Sequence 2125, Ap
33	36.6	9.9	447	1	PCT-US02-10421-2349	Sequence 2349, Ap
34	36.6	9.9	447	7	US-10-112-699-2349	Sequence 2349, Ap
35	36.6	9.9	691	7	US-10-162-749-26	Sequence 26, Appl
36	36.6	9.9	695	7	US-10-143-788-50	Sequence 50, Appl
37	36.6	9.9	695	7	US-10-162-749-52	Sequence 52, Appl
38	36.6	9.9	700	7	US-10-143-788-53	Sequence 53, Appl
39	36.6	9.9	700	7	US-10-162-749-67	Sequence 67, Appl
40	36.6	9.9	746	7	US-10-162-749-11	Sequence 11, Appl
41	36.2	9.8	195	5	US-09-442-366A-1076	Sequence 1076, Ap
42	36.2	9.8	305	1	PCT-US02-10421-1244	Sequence 1244, Ap
43	36.2	9.8	305	7	US-10-112-699-1244	Sequence 1244, Ap
44	35	9.5	439	1	PCT-US02-10421-2521	Sequence 2521, Ap
45	35	9.5	439	7	US-10-112-699-2521	Sequence 2521, Ap

## ALIGNMENTS

RESULT 1  
US-10-162-749-16  
Sequence 16, Application US/10162749  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PTZ67C1N  
CURRENT APPLICATION NUMBER: US/10/162,749  
CURRENT FILING DATE: 2002-06-06  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 184  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 16  
LENGTH: 603  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (584)  
OTHER INFORMATION: n equals a,t,g, or c  
US-10-162-749-16

Query Match	93.7%	Score 344.8	DB 7	Length 603
Best Local Similarity	96.8%	Pred. No. 5.5e-100		
Matches	366	Conservative	0	Mismatches 2; Indels 10; Gaps 1
Oy	1	gaagagacgtgtaagtgcggtgcagttttcaactgaccttgacgcagaaacttcagcc	60	
Db	177	gaagagacgtgtaagtgcggtgcagttttcaactgaccttgacgcagaaacttcagcc	236	
Oy	51	atgaaggttaacaggcattcttctcagtcgcttgcccttgagtcgtatctctggttaac	120	
Db	237	atgaaggttaacaggcattcttctcagtcgcttgcccttgagtcgtatctctggttaac	296	
Oy	121	actgagtcgtactccctggagagagggccaaatgttacaatgaacttaagtgcgacc	180	
Db	297	actgagtcgtactccctggagagagggccaaatgttacaatgaacttaagtgcgacc	356	
Oy	181	aagatatgacctgtctgtggaactgagtgaaatacttaccatgaatgcgtgta	240	
Db	357	aagatatgacctgtctgtggaactgagtgaaatacttaccatgaatgcgtgta	416	
Oy	261	tggtttgaagtcgaaacgcagacttctctcattcaaaaatctggccttctgta	300	
Db	417	tggtttgaagtcgaaacgcagacttctctcattcaaaaatctggccttctgta	476	
Oy	301	gaaccaagttttgaaatccctcaggtcacccgcgagggc-----tattgttgaa	350	

Db 477 gaaccagggtttgaatcccatcaggtcaccgcgaggcctgactgagccttattgttgaa 536  
Qy 351 taaatgatctgaatc 368  
Db 537 taaatgatctgaatc 554

RESULT 2  
US-10-162-749-57/c  
: Sequence 57, Application US/10162749  
: GENERAL INFORMATION:  
: APPLICANT: Rosen et al.  
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
: FILE REFERENCE: PTZ67CIN  
: CURRENT APPLICATION NUMBER: US/10/162,749  
: CURRENT FILING DATE: 2002-06-06  
: Prior Application removed - See File Wrapper or Palm  
: NUMBER OF SEQ ID NOS: 184  
: SOFTWARE: PatentIn Ver. 2.0  
: SEQ ID NO 57  
: LENGTH: 607  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
: FEATURE:  
: NAME/KEY: misc\_feature  
: LOCATION: (23)  
: OTHER INFORMATION: n equals a,t,g, or c  
: FEATURE:  
: NAME/KEY: misc\_feature  
: LOCATION: (25)  
: OTHER INFORMATION: n equals a,t,g, or c  
: FEATURE:  
: NAME/KEY: misc\_feature  
: LOCATION: (35)  
: OTHER INFORMATION: n equals a,t,g, or c  
US-10-162-749-57

Query Match 93.7% Score 344.8; DB 7; Length 607;  
Best Local Similarity 96.8%; Pred. No. 5, 5e-100;  
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

Qy 1 gaagagcgtggaagtcgagtcagttttcaactgacctgtgacacgaactcagcc 60  
Db 429 GAAGAGCGTGTGAAGTCGGTGCAGTTCACACTGACCTTGGACGACAGAACTTCAGCC 370  
Qy 61 atgaagtaacagacatcttcttctcagtcgcttgcctgttgagtcctatctggtaac 120  
Db 369 ATGAAGGTAACAGGCACTTCTTCTCAGTCGCTTGGCCCTGTTGAGCTAICTGTTAAC 310  
Qy 121 actgagctgactcctcctgggaagagagagccaaatgttacaatgaacttaattggtgacc 180  
Db 309 ACTGGAGCTGACTCCTCGGAAGAGAGAGGCCAAATGTTACAAATGAACITTAATGGATGCACC 250  
Qy 181 aagatatagccttctgtgagcactgagaaacttatcccaatgaatgcgttta 240  
Db 249 AAGATATATGACCCCTGTCTGTGGGACTGATGGAATACTTATCCCAATGAATGCGTTA 190  
Qy 241 tgtttgaaggtcggaacgcgcagactctctatctcattcaaaaatctgggacctgtga 300  
Db 189 TGTTTGAAAATCGGAACGCCAGACTTCTATCTCATTAACAAAATCTGGGCCCTTGCTGA 130  
Qy 301 gaaccaaggttttgaatcccatcaggtcaccgcgagcc-----tatgttgaa 350  
Db 129 GAACCAAGGTTTGAATCCCAATCAGGTACCGGACCGGAGCCCTTATGTTGAA 70  
Qy 351 taaatgatctgaatc 368  
Db 69 TAAATGATCTGAATATC 52

RESULT 3

US-60-382-445-457  
: Sequence 457, Application US/60382445  
: GENERAL INFORMATION:  
: APPLICANT: Algate, Paul A.  
: APPLICANT: Mannion, Jane  
: APPLICANT: Gaiger, Alexander  
: APPLICANT: Gordon, Brian  
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
: FILE REFERENCE: 210121.573PI  
: CURRENT APPLICATION NUMBER: US/60/382,445  
: CURRENT FILING DATE: 2002-05-20  
: NUMBER OF SEQ ID NOS: 2900  
: SOFTWARE: FastSeq for Windows Version 4.0  
: SEQ ID NO 457  
: LENGTH: 214  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
US-60-382-445-457

Query Match 57.3% Score 210.8; DB 8; Length 214;  
Best Local Similarity 99.1%; Pred. No. 2, 4e-57;  
Matches 212; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 125 gagctgactcctcctgggaagagagagccaaatgttacaatgaacttaattggtgacccaaga 184  
Db 1 gagctgactcctcctgggaagagagagccaaatgttacaatgaacttaattggtgacccaaga 60  
Qy 125 tatatgacctgtctgtgggactgatggaatacttaccatgaatgcgtgttatgt 244  
Db 61 tatatgacctgtctgtgggactgatggaatacttaccatgaatgcgtgttatgt 120  
Qy 245 ttgaaggtcggaacgcgcagactctctcattcaaaaatctgggacctgtgagaac 304  
Db 121 ttgaaatcggaacgcgcagactctctcattcaaaaatctgggacctgtgagaac 180  
Qy 305 taaggtttgaaatcccatcaggtcaccgcgagg 338  
Db 191 taaggtttgaaatcccatcaggtcaccgcgagg 214

RESULT 4  
US-09-454-226A-557  
: Sequence 557, Application US/09454226A  
: GENERAL INFORMATION:  
: APPLICANT: Chenchik, Alex  
: APPLICANT: Lukashev, Matvey  
: TITLE OF INVENTION: Rat Array  
: FILE REFERENCE: CLON-006CIP12  
: CURRENT APPLICATION NUMBER: US/09/454,226A  
: CURRENT FILING DATE: 2002-05-07  
: NUMBER OF SEQ ID NOS: 1186  
: SOFTWARE: FastSeq for Windows Version 4.0  
: SEQ ID NO 557  
: LENGTH: 180  
: TYPE: DNA  
: ORGANISM: Artificial Sequence  
: FEATURE:  
: OTHER INFORMATION: Synthesized gene fragment  
US-09-454-226A-557

Query Match 24.1% Score 88.8; DB 5; Length 180;  
Best Local Similarity 68.3%; Pred. No. 2, 1e-18;  
Matches 123; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 101 tgttgagctctatctgttaacactgagactcctccctgggaagagagccaaatgttaca 160  
Db 1 tgcctcagtttagcaggtaacccctccagctgaggtgaatggaataacacccctaatgtgccta 60  
Qy 101 atgaacttaattggtgacccaagatatatgacctgtctgtgggactgtgtggaatactt 220



Db 160 TGTCCAGATGTCACCTGCTGCGGCACTGATGGCTCACATATACGAATGATGC 101  
 Oy 235 gtgttatgtttgaaggtcggaacgcccagacttctatctcattcaaaaatctgggcct 294  
 Db 100 CAGCTCTGCTTGGCCGGATATAAAACCAACAGGACATCCAGATCATGAAGAATGGCAAA 41  
 Oy 295 tgcctgagaacaaagg 309  
 Db 40 TGCTGATCCACAGG 26

RESULT 8  
 PCT-US01-43704-2330  
 : Sequence 2330, Application PC/TUS0143704  
 : GENERAL INFORMATION:  
 : APPLICANT: Corixa Corporation  
 : APPLICANT: Stolck, John A.  
 : APPLICANT: Xu, Jiangchun  
 : APPLICANT: Chenault, Ruth A.  
 : APPLICANT: Meagher, Madelein Joy  
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 : TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER  
 : FILE REFERENCE: 210121.561PC  
 : CURRENT APPLICATION NUMBER: PCT/US01/43704  
 : CURRENT FILING DATE: 2001-11-19  
 : NUMBER OF SEQ ID NOS: 2606  
 : SOFTWARE: Corixa Invention Disclosure Database  
 : SEQ ID NO 2330  
 : LENGTH: 323  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 PCT-US01-43704-2330

Query Match 11.5%; Score 42.2; DB 1; Length 323;  
 Best Local Similarity 57.0%; Pred. No. 0.0019;  
 Matches 77; Conservative 0; Mismatches 58; Indels 0; Gaps 0;  
 Oy 175 tgcacaaagatatgacctgtctgtgggactgtggaatacttctcattcaaaaatctgggcct 234  
 Db 164 tgtccagatgtccaaacctgctgctggcactgtggtctcacatatacgaatgaatgc 223  
 Oy 235 gtgttatgtttgaaggtcggaacgcccagacttctatctcattcaaaaatctgggcct 294  
 Db 224 cagctctgtctggccggataaaacacacagacatccagatcatgaagatggcaaa 283  
 Oy 295 tgcctgagaacaaagg 309  
 Db 284 tgctgacccacagg 298

RESULT 9  
 US-09-671-483-2  
 : Sequence 2, Application US/09671483  
 : GENERAL INFORMATION:  
 : APPLICANT: Kato, Seishi  
 : APPLICANT: Yamaguchi, Tomoko  
 : APPLICANT: Sekine, Shingo  
 : APPLICANT: Kamata, Kouju  
 : TITLE OF INVENTION: HUMAN PEC-60-LIKE PROTEIN AND DNA ENCODING THIS  
 : TITLE OF INVENTION: PROTEIN  
 : FILE REFERENCE: 6700PCT-US  
 : CURRENT APPLICATION NUMBER: US/09/671.483  
 : CURRENT FILING DATE: 2000-09-27  
 : PRIOR APPLICATION NUMBER: US/09/065,019  
 : PRIOR FILING DATE: 1998-04-17  
 : NUMBER OF SEQ ID NOS: 3  
 : SOFTWARE: Patent In Ver. 2.0  
 : SEQ ID NO 2  
 : LENGTH: 398  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens

US-09-671-483-2  
 Query Match 11.5%; Score 42.2; DB 5; Length 398;  
 Best Local Similarity 57.0%; Pred. No. 0.002;  
 Matches 77; Conservative 0; Mismatches 58; Indels 0; Gaps 0;  
 Oy 175 tgcacaaagatatgacctgtctgtgggactgtggaatacttctcattcaaaaatctgggcct 234  
 Db 178 tgtccagatgtccaaacctgctgctggcactgtggtctcacatatacgaatgaatgc 237  
 Oy 235 gtgttatgtttgaaggtcggaacgcccagacttctatctcattcaaaaatctgggcct 294  
 Db 238 cagctctgtctggccggataaaacacacagacatccagatcatgaagatggcaaa 297  
 Oy 295 tgcctgagaacaaagg 309  
 Db 298 tgctgacccacagg 312

RESULT 10  
 US-10-106-698-1897  
 : Sequence 1897, Application US/10106698  
 : GENERAL INFORMATION:  
 : APPLICANT: Ruben et al.  
 : TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept  
 : FILE REFERENCE: P400SP1  
 : CURRENT APPLICATION NUMBER: US/10/106,698  
 : CURRENT FILING DATE: 2002-03-27  
 : PRIOR APPLICATION NUMBER: PCT/US00/26524  
 : PRIOR FILING DATE: 2000-09-28  
 : PRIOR APPLICATION NUMBER: US 60/157,137  
 : PRIOR FILING DATE: 1999-09-29  
 : PRIOR APPLICATION NUMBER: US 60/163,280  
 : PRIOR FILING DATE: 1999-11-03  
 : NUMBER OF SEQ ID NOS: 8564  
 : SOFTWARE: Patent In Ver. 3.0  
 : SEQ ID NO 1897  
 : LENGTH: 433  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 US-10-106-698-1897

Query Match 11.5%; Score 42.2; DB 7; Length 433;  
 Best Local Similarity 57.0%; Pred. No. 0.0021;  
 Matches 77; Conservative 0; Mismatches 58; Indels 0; Gaps 0;  
 Oy 175 tgcacaaagatatgacctgtctgtgggactgtggaatacttctcattcaaaaatctgggcct 234  
 Db 194 tgtccagatgtccaaacctgctgctggcactgtggtctcacatatacgaatgaatgc 253  
 Oy 235 gtgttatgtttgaaggtcggaacgcccagacttctatctcattcaaaaatctgggcct 294  
 Db 254 cagctctgtctggccggataaaacacacagacatccagatcatgaagatggcaaa 313  
 Oy 295 tgcctgagaacaaagg 309  
 Db 314 tgctgacccacagg 328

RESULT 11  
 PCT-US02-10824-19  
 : Sequence 19, Application PC/TUS0210824  
 : GENERAL INFORMATION:  
 : APPLICANT: Origene Technologies  
 : TITLE OF INVENTION: Prostate Cancer Expression Profiles  
 : FILE REFERENCE: 9U 206 PCT  
 : CURRENT APPLICATION NUMBER: PCT/US02/10824  
 : CURRENT FILING DATE: 2002-04-08  
 : PRIOR APPLICATION NUMBER: US 60/281,732  
 : PRIOR FILING DATE: 2001-04-06  
 : PRIOR APPLICATION NUMBER: US 60/281,731

```

; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 3528
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-10824-19

Query Match      11.0%; Score 40.6; DB 1; Length 3528;
Best Local Similarity 57.5%; Pred. No. 0.013;
Matches 73; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 170 atggatgcccaagaatataatgacctgtctgtggactgatggaataacttaccctgaatg 229
    || || || || || || || || || || || || || || || || || || || || ||
Db 3054 atctttgtccaaaggatttaaagcctgtctgtggactgatggaataacttaccctgaatc 3113
    || || || || || || || || || || || || || || || || || || || || ||

QY 230 aatgcgtgtattgtttgaagtcggaacgccagacttctatctctcattcaaaaatctg 289
    || || || || || || || || || || || || || || || || || || || || ||
Db 3114 ctgcatgctctgtcatgaaacccgtatagcgaacaaatacacacatccgcagtagacag 3173
    || || || || || || || || || || || || || || || || || || || || ||

QY 290 ggccttg 296
    || || ||
Db 3174 ggaagtg 3180

RESULT 12
US-09-926-820-7
; Sequence 7, Application US/09926820
; GENERAL INFORMATION:
; APPLICANT: FORSSMANN, WOLF-GEORG
; APPLICANT: MAGERT, HANS-JURGEN
; APPLICANT: STANDKER, LUDGER
; APPLICANT: KREUTZMANN, PETER
; TITLE OF INVENTION: SERINE PROTEASE INHIBITORS
; FILE REFERENCE: 10496-P67431US0
; CURRENT APPLICATION NUMBER: US/09/926,820
; PRIOR FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: PCT/EP99/04331
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 3532
; TYPE: DNA
; ORGANISM: mammalian
; NAME/KEY: CDS
; LOCATION: (44)...(3235)
US-09-926-820-7

Query Match      11.0%; Score 40.6; DB 5; Length 3532;
Best Local Similarity 57.5%; Pred. No. 0.013;
Matches 73; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 170 atggatgcccaagaatataatgacctgtctgtggactgatggaataacttaccctgaatg 229
    || || || || || || || || || || || || || || || || || || || || ||
Db 3054 atctttgtccaaaggatttaaagcctgtctgtggactgatggaataacttaccctgaatc 3113
    || || || || || || || || || || || || || || || || || || || || ||

QY 230 aatgcgtgtattgtttgaagtcggaacgccagacttctatctctcattcaaaaatctg 289
    || || || || || || || || || || || || || || || || || || || || ||
Db 3114 ctgcatgctctgtcatgaaacccgtatagcgaacaaatacacacatccgcagtagacag 3173
    || || || || || || || || || || || || || || || || || || || || ||

QY 290 ggccttg 296
    || || ||
Db 3174 ggaagtg 3180

RESULT 13
US-09-926-820-12
; Sequence 12, Application US/09926820

```

```

; GENERAL INFORMATION:
; APPLICANT: FORSSMANN, WOLF-GEORG
; APPLICANT: MAGERT, HANS-JURGEN
; APPLICANT: STANDKER, LUDGER
; APPLICANT: KREUTZMANN, PETER
; TITLE OF INVENTION: SERINE PROTEASE INHIBITORS
; FILE REFERENCE: 10496-P67431US0
; CURRENT APPLICATION NUMBER: US/09/926,820
; PRIOR FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: PCT/EP99/04331
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 156
; TYPE: DNA
; ORGANISM: mammalian
US-09-926-820-12

Query Match      11.0%; Score 40.4; DB 5; Length 156;
Best Local Similarity 58.2%; Pred. No. 0.0056;
Matches 71; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 170 atggatgcccaagaatataatgacctgtctgtggactgatggaataacttaccctgaatg 229
    || || || || || || || || || || || || || || || || || || || || ||
Db 32 atctttgtccaaaggatttaaagcctgtctgtggactgatggaataacttaccctgaatc 91
    || || || || || || || || || || || || || || || || || || || || ||

QY 230 aatgcgtgtattgtttgaagtcggaacgccagacttctatctctcattcaaaaatctg 289
    || || || || || || || || || || || || || || || || || || || || ||
Db 52 ttgcatgctctgtcatgaaacccgtatagcgaacaaatacacacatccgcagtagacag 151
    || || || || || || || || || || || || || || || || || || || || ||

QY 290 qg 291
    || ||
Db 152 qg 153

RESULT 14
US-09-529-469A-8198/c
; Sequence 8198, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORI
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8198

```

; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11819  
; LENGTH: 1022  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-629-469A-11819

Query Match 10.7%; Score 39.2; DB 5; Length 1022;  
Best Local Similarity 58.6%; Pred. No. 0.025;  
Matches 68; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Oy 175 tgcaccaagatatatgaccctgtctgtgggactgatggaataacttaccatgaatgc 234  
Db 688 tgcaccaatgtgacagcaccctgttgcctcaaatggccacacttccagaatgagt 747  
Oy 235 gtgttatgttttgaaggcggaacgcagacttctatccctcattcaaaaatctgg 290  
Db 748 tcttttgtgtgaaacagagggaatttcattatcgataaaaaatttgaaaatatgg 803

Search completed: July 29, 2002, 19:27:21  
Job time: 7612 sec

LENGTH: 562  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (484)..(484)  
OTHER INFORMATION: "n" may be a, t, c, g, other or unknown  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (505)..(505)  
OTHER INFORMATION: "n" may be a, t, c, g, other or unknown  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (552)..(552)  
OTHER INFORMATION: "n" may be a, t, c, g, other or unknown  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (560)..(560)  
OTHER INFORMATION: "n" may be a, t, c, g, other or unknown  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (562)..(562)  
OTHER INFORMATION: "n" may be a, t, c, g, other or unknown  
US-09-629-469A-8198

Query Match 10.7%; Score 39.2; DB 5; Length 562;  
Best Local Similarity 58.6%; Pred. No. 0.021;  
Matches 68; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Oy 175 tgcaccaagatatatgaccctgtctgtgggactgatggaataacttaccatgaatgc 234  
Db 335 TGCCCCAATGTGACAGCACCTGTTGTGCTCAAAATGGCCACACTTCCAGAAATGAGTGT 275  
Oy 235 gtgttatgttttgaaggcggaacgcagacttctatccctcattcaaaaatctgg 290  
Db 275 TTCCTTTGTGTGAACAGAGGGAATTCATTATCGTATAAAATTTGAAAAATATGG 220

RESULT 15  
US-09-629-469A-11819  
; Sequence 11819, Application US/09629469A  
; GENERAL INFORMATION:  
; APPLICANT: OTA, TOSHIO  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: NISHIKAWA, TETSUO  
; APPLICANT: HAYASHI, KOJI  
; APPLICANT: SAITO, KAORU  
; APPLICANT: YAMAMOTO, JUNICHI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: WAKAMATSU, AI  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: OTSUKI, TETSUJI  
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE  
; FILE REFERENCE: 084335/0123  
; CURRENT APPLICATION NUMBER: US/09/629,469A  
; PRIOR FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: JP 1999-248036  
; PRIOR FILING DATE: 1999-07-29  
; PRIOR APPLICATION NUMBER: JP 1999-300253  
; PRIOR FILING DATE: 1999-08-27  
; PRIOR APPLICATION NUMBER: JP 2000-118776  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: JP 2000-183767  
; PRIOR FILING DATE: 2000-05-02  
; PRIOR APPLICATION NUMBER: JP 2000-241899  
; PRIOR FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: 60/159,590  
; PRIOR FILING DATE: 1995-10-18  
; PRIOR APPLICATION NUMBER: 60/183,322  
; PRIOR FILING DATE: 2000-02-17  
; NUMBER OF SEQ ID NOS: 19025

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 17:54:40 : Search time 100.77 seconds  
(without alignments)  
897.024 Million cell updates/sec

Title: US-09-880-107-3847

Perfect score: 368  
Sequence: 1 gaagagacgtgtaagtgcg.....aataaatgtatctgaatc 368

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.NA.\*

- 1: /cgn2.6/ptodata/2/ina/5A\_COMB.seq.\*
- 2: /cgn2.6/ptodata/2/ina/5B\_COMB.seq.\*
- 3: /cgn2.6/ptodata/2/ina/6A\_COMB.seq.\*
- 4: /cgn2.6/ptodata/2/ina/6B\_COMB.seq.\*
- 5: /cgn2.6/ptodata/2/ina/PCTUS\_COMB.seq.\*
- 6: /cgn2.6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	211.4	57.4	265	US-09-221-298-55	Sequence 55, Appl
2	162.2	44.1	420	US-09-221-298-56	Sequence 56, Appl
3	162	44.0	302	US-08-235-515A-26	Sequence 26, Appl
4	109	29.6	458	US-08-880-829-21	Sequence 21, Appl
5	42.2	11.5	388	US-08-744-670-2	Sequence 2, Appl
6	42.2	11.5	388	US-09-149-933-2	Sequence 2, Appl
7	42.2	11.5	398	US-09-065-019-2	Sequence 2, Appl
8	38.2	10.4	180	US-09-065-019-1	Sequence 1, Appl
9	36	9.8	1565	US-08-468-846-1	Sequence 1, Appl
10	36	9.8	1565	US-08-915-096A-1	Sequence 1, Appl
11	35.4	9.6	126	US-08-586-676E-4	Sequence 4, Appl
12	35.4	9.6	129	US-08-586-676E-5	Sequence 5, Appl
13	35.4	9.6	138	US-08-586-676E-6	Sequence 6, Appl
14	34.6	9.4	149	US-08-586-676E-7	Sequence 7, Appl
15	33.8	9.2	149	US-08-586-676E-8	Sequence 8, Appl
16	33.4	9.1	4800	PCT-US94-07779-1	Sequence 1, Appl
17	33	9.0	1282	US-08-211-942-16	Sequence 16, Appl
18	32.8	8.9	309	US-08-839-709-2	Sequence 2, Appl
19	32.8	8.9	309	US-09-204-859-2	Sequence 2, Appl
20	31	8.4	2728	US-09-025-769B-299	Sequence 299, App
21	30.8	8.4	1541	US-08-446-924-1	Sequence 1, Appl
22	30.8	8.4	1541	US-08-798-665-1	Sequence 1, Appl
23	30.8	8.4	1541	US-08-982-987A-1	Sequence 1, Appl
24	30.6	8.3	1896	US-09-345-468-24	Sequence 24, Appl
25	30.6	8.3	2170	US-09-345-468-11	Sequence 11, Appl
26	30	8.2	3195	US-08-951-648-5	Sequence 5, Appl
27	30	8.2	3195	US-09-174-437-5	Sequence 5, Appl

28 29 7.9 908 4 US-09-439-313-350 Sequence 350, Appl  
29 29 7.9 1122 2 US-08-820-170A-26 Sequence 26, Appl  
30 29 7.9 1122 3 US-09-055-699-26 Sequence 26, Appl  
31 29 7.9 1122 4 US-09-273-565-26 Sequence 26, Appl  
32 29 7.9 1122 4 US-09-565-538-26 Sequence 26, Appl  
33 29 7.9 1346 3 US-08-949-202-3 Sequence 3, Appl  
34 29 7.9 1346 4 US-09-418-175-3 Sequence 3, Appl  
35 29 7.9 1695 5 PCT-US95-06385-1 Sequence 1, Appl  
36 29 7.9 1721 2 US-08-820-170A-27 Sequence 27, Appl  
37 29 7.9 1721 3 US-09-055-699-27 Sequence 27, Appl  
38 29 7.9 1721 4 US-09-273-565-27 Sequence 27, Appl  
39 29 7.9 1721 4 US-09-565-538-27 Sequence 27, Appl  
40 28.8 7.8 1057 3 US-09-188-930-18 Sequence 18, Appl  
41 28.8 7.8 2004 3 US-09-188-930-230 Sequence 230, App  
42 28.8 7.8 12839 3 US-09-125-287-1 Sequence 1, Appl  
43 28.6 7.8 6765 2 US-08-677-010-2 Sequence 2, Appl  
44 28.6 7.8 6765 2 US-08-790-519-2 Sequence 2, Appl  
45 28.6 7.8 9581 2 US-08-677-010-1 Sequence 1, Appl

#### ALIGNMENTS

RESULT 1  
US-09-221-298-55  
: Sequence 55, Application US/09221298  
: Patent No. 6284241  
: GENERAL INFORMATION:  
: APPLICANT: Xu, Jiangchun  
: TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND DIAGNOSIS  
: TITLE OF INVENTION: OF COLON CANCER  
: FILE REFERENCE: 210121.471  
: CURRENT APPLICATION NUMBER: US/09/221,298  
: CURRENT FILING DATE: 1998-12-23  
: NUMBER OF SEQ ID NOS: 112  
: SOFTWARE: FastSeq for Windows Version 3.0  
: SEQ ID NO 55  
: LENGTH: 265  
: TYPE: DNA  
: ORGANISM: Human  
: FEATURE:  
: NAME/KEY: modified\_base  
: LOCATION: (25)  
: OTHER INFORMATION: Where n is a, c, g or t  
: FEATURE:  
: NAME/KEY: modified\_base  
: LOCATION: (223)  
: OTHER INFORMATION: Where n is a, c, g or t  
: FEATURE:  
: NAME/KEY: modified\_base  
: LOCATION: (241)  
: OTHER INFORMATION: Where n is a, c, g or t  
: FEATURE:  
: NAME/KEY: modified\_base  
: LOCATION: (254)  
: OTHER INFORMATION: Where n is a, c, g or t  
: FEATURE:  
: NAME/KEY: modified\_base  
: LOCATION: (259)  
: OTHER INFORMATION: Where n is a, c, g or t  
US-09-221-298-55

Query Match 57.4%; Score 211.4; DB 4; Length 265;  
Best Local Similarity 95.6%; Pred. No. 4.2e-64;  
Matches 237; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

Oy 76 atctttcttcagtccttgccctgttgagctctatctgtaacactggagctgactcc 135

Db 1 atctttcttcagtccttgccctgttgagctctatctgtaacactggagctgactcc 60

Oy 136 ctgggaagagagcgcacaaatgttacaatgaacttaagtgaacacacgaatatatgacct 195





```

; ; OTHER INFORMATION: /note= "j-x1-y-z(psti), insert in plasmid pM474,
; ; OTHER INFORMATION: Figure 5"
; ;
; ; FEATURE:
; ;   NAME/KEY: sig_peptide
; ;   LOCATION: 27..89
; ;
; ; FEATURE:
; ;   NAME/KEY: mat_peptide
; ;   LOCATION: 90..293
; ;
; ; FEATURE:
; ;   NAME/KEY: CDS
; ;   LOCATION: 27..293
; ;
US-08-235-515A-26

```

Query Match	44.0%	Score 162	DB 2	Length 302
Best Local Similarity	91.9%	Pred. No. 6.5e-47		
Matches 171	Conservative	0	Mismatches 15	Indels 0
Gaps				
120	cactggagctgactccctgggaagagacaaatgtttacaatgaacttaatgaatgcac	179		
116	CTCGGGATGGACTCCCTAGGTGCGGAGCCAAATGTTACAATGAACITTAATGATGCAC	175		
180	caagatatatgacctgtctctgggaactgatggaataacttatcccaatgaatgcgtgtt	239		
176	CAAGATATATGACCCCTGCTGTGGGACTGATGGAAATACTTATCCCAATGAATGCGTGT	235		
240	atgttttgaagatcggaagaccagactcttatctctcatcctcaaaaatctgggacctgtcg	299		
236	ATGTTTTGAAAATCGGAAACGCCAGACATCGATCTCATTCAAAAATCTGGGCCCTTGCTG	295		
300	agaacc 305			
296	AGGATC 301			

RESULT 4  
US-08-880-829-21  
: Sequence 21, Application US/08880829  
: Patent No. 592559  
: GENERAL INFORMATION:  
: APPLICANT: Collins, John  
: APPLICANT: Roettgen, Peter  
: TITLE OF INVENTION: A Collection of Phagemids, A  
: TITLE OF INVENTION: Cells Carrying The Phagemids, A  
: TITLE OF INVENTION: Collection of Phagemid Particles  
: TITLE OF INVENTION: Produced From Said Collection  
: TITLE OF INVENTION: And Phagemid Particles  
: NUMBER OF SEQUENCES: 21  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Joseph T. Eisele  
: ADDRESSEE: Kane, Dalsimer, Sullivan, Kurucz,  
: ADDRESSEE: Levy, Eisele and Richard  
: STREET: 711 Third Avenue  
: CITY: New York  
: STATE: New York  
: COUNTRY: U.S.A.  
: ZIP: 10017-4059  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: 3-1/2" DISKETTE  
: COMPUTER: IBM-XT COMPATIBLE  
: OPERATING SYSTEM: DOS 3.3:  
: SOFTWARE: WORDPERFECT 5.0  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/880, 829  
: FILING DATE: 23-JUN-1997  
: CLASSIFICATION: 4.35  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: 08/458, 668  
: FILING DATE: 06/02/95  
: APPLICATION NUMBER: German EP 94 108 689.4  
: FILING DATE: 06/07/94  
: ATTORNEY/AGENT INFORMATION:  
:

```

NAME: EISELE, JOSEPH T.
REGISTRATION NUMBER: 25,331
REFERENCE/DOCKET NUMBER: 2727-77
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 687-6000
TELEFAX: (212) 682-3485
TELEX: (212) 426767
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 458 base pairs
TYPE: nucleic acid
STRANDEDNESS: single strand
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
;
US-08-890-929-21

Query Match          29.6%; Score 109; DB 2: Length 458;
Best Local Similarity 69.5%; Pred. No. 2.1e-28;
Matches 140 Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Qy    68   agtgccttgccctgttgagtcatactcgtgtaaacactggagctgaactcccctggggaagagag 147
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    133  ATTGCAGTGGCATTGCGTGCTGCCGTACCCGACGGCAGCCGCACTCTCTGGGTGCTGA 192
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy    143  gccaaatgtcacatgaacctaattgatgcaccacaagatatatgacctgtctctgtggact 207
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    193  GCTAAATGCTACAACGAGCTCAACGGTTGCACTAAGATCTACGACCCGGTTTCGGGTACC 252
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy    208  gatggaataacttatccccaatgaatcggtttaagtattttgaaggtcgaaaacgccagact 267
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    253  GACGGCAACACTTACCCGCAAGAAATGGGTTCTGTGCTTCGAAAACCGTTAACACGTACACT 312
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy    268  ttatctctcatcaaaaaatctgggccttgtcta 300
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
b     313  TCTATCTCTGATCCAGAATACTGGTCCGTGCTTA 345

```

```

RESULT      5
US-08-744-570-2
: Sequence 2, Application US/08744670
: Patent No. 5858710
: GENERAL INFORMATION:
: APPLICANT: Bandman, Olga
: APPLICANT: Goli, Surya K.
: APPLICANT: Murry, Lynn E.
: TITLE OF INVENTION: TUMOR-ASSOCIATED KAZAL INHIBITOR
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: US
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS

```

SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/744,670  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0155 US  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 388 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: SINTTUT01  
CLONE: 1539065  
US-08-744-670-2

Query Match 11.5% Score 42.2; DB 2; Length 388;  
Best Local Similarity 57.0%; Pred. No. 3.1e-05;  
Matches 77; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Oy 175 tgcaccaagatatagccctgtctgtgggactgatgaaatacttaccatgaatgc 234  
Db 168 TGTTCACAGATGTCACCTGCTGCGGACTGATGGCTCACATATACGAATGAATGC 227

Oy 235 gtgttatgtttgaagtcggaacgccagactctctctcattcctcctcaaaaatctggcct 294  
Db 228 CAGCTCTGCTTGGCCGGATAAAACCAACAGGACATCCAGATCATGAAAGATGGCAAA 287

Oy 295 tgcagagaaccaagg 309  
Db 288 TGCTGATCCACAGG 302

RESULT 6  
US-09-149-933-2  
Sequence 2, Application US/09149933  
Patent No. 5958699  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Goli, Surya K.  
APPLICANT: Murry, Lynn E.  
TITLE OF INVENTION: TUMOR-ASSOCIATED KAZAL INHIBITOR  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/149,933  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0155 US  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 388 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: SINTTUT01  
CLONE: 1539065  
US-09-149-933-2

Query Match 11.5% Score 42.2; DB 2; Length 388;  
Best Local Similarity 57.0%; Pred. No. 3.1e-05;  
Matches 77; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Oy 175 tgcaccaagatatagccctgtctgtgggactgatgaaatacttaccatgaatgc 234  
Db 168 TGTTCACAGATGTCACCTGCTGCGGACTGATGGCTCACATATACGAATGAATGC 227

Oy 235 gtgttatgtttgaagtcggaacgccagactctctctcattcctcctcaaaaatctggcct 294  
Db 228 CAGCTCTGCTTGGCCGGATAAAACCAACAGGACATCCAGATCATGAAAGATGGCAAA 287

Oy 295 tgcagagaaccaagg 309  
Db 288 TGCTGATCCACAGG 302

RESULT 7  
US-09-065-019-2  
Sequence 2, Application US/09065019  
Patent No. 6280968  
GENERAL INFORMATION:  
APPLICANT: Kato, Seishi  
APPLICANT: Yamaguchi, Tomoko  
APPLICANT: Sekine, Shingo  
APPLICANT: Kamata, Kouju  
TITLE OF INVENTION: HUMAN PEC-60-LIKE PROTEIN AND DNA ENCODING THIS PROTEIN  
FILE REFERENCE: 6700PCT-US  
CURRENT APPLICATION NUMBER: US/09/065,019  
CURRENT FILING DATE: 1998-04-17  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 398  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-065-019-2

Query Match 11.5% Score 42.2; DB 4; Length 398;  
Best Local Similarity 57.0%; Pred. No. 3.2e-05;  
Matches 77; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Oy 175 tgcaccaagatatagccctgtctgtgggactgatgaaatacttaccatgaatgc 234  
Db 178 tgttccacagatgtcccaacctggctcgtcggcactgatggctcacatatacgaatgaatgc 237

Oy 235 gtgttatgtttgaagtcggaacgccagactctctcattcctcctcaaaaatctggcct 294  
Db 238 cagctcgtctggcccgataaaacacacagagacatccagatcatgaagaatggcaaa 297

Oy 295 tgcagagaaccaagg 309  
Db 298 tgcagatccccacagg 312

```

RESULT      8
US-09-065-019-1
: Sequence 1, Application US/09065019
: Patent No. 6280988
: GENERAL INFORMATION:
: APPLICANT: Kato, Seishi
: APPLICANT: Yamaguchi, Tomoko
: APPLICANT: Sekine, Shingo
: APPLICANT: Kamata, Kouju
: TITLE OF INVENTION: HUMAN PEC-60-LIKE PROTEIN AND DNA ENCODING THIS PROTEIN
: FILE REFERENCE: 6700PCT-US
: CURRENT APPLICATION NUMBER: US/09/065,019
: CURRENT FILING DATE: 1998-04-17
: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 1
: LENGTH: 180
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-065-019-1

```

	Query Match	10.4%	Score 38.2;	DB 4;	Length 180;
	Best Local Similarity	56.9%;	Pred. No. 0.00054;		
Matches	70; Conservative	0;	Mismatches 53;	Indels	Gaps 0;
Oy	175	tgaccacaagatatgacccctgtgtggagactgatgaataactatcccagaatgc	234		
Db	58	tgttccagatgccaaacctggtctcgcgcactgatggtcacatatacgaatgaatgc	117		
Oy	235	gttgtatgtttgaaggctcggaacgccagactctctctcatctcaaaaatctggcct	294		
Db	118	cagctctgttggccggatataaaaaaaccaagacatccagatcatgaagaatggcaaa	177		
Oy	295	tgc	297		
Db	178	tgc	180		

RESULT 9  
US-08-468-846-1  
: Sequence 1, Application US/08468846  
: Patent No. 6074839  
: GENERAL INFORMATION:  
: APPLICANT: Meissner, Paul  
: APPLICANT: Fuldner, Rebecca  
: APPLICANT: Fei-wei, Ying  
: APPLICANT: Adams, Mark  
: TITLE OF INVENTION: TRANSFORMING GROWTH FACTOR ALPHA HI  
: NUMBER OF SEQUENCES: 15  
: CORRESPONDENCE ADDRESSES:  
: ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,  
: ADDRESSEE: STUART & OLSTEIN  
: STREET: 6 Becker Farm Road  
: CITY: Roseland  
: STATE: New Jersey  
: COUNTRY: USA  
: ZIP: 07068  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/468,846  
: FILING DATE: 06-JUN-1995  
: CLASSIFICATION: 514  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/208,008  
: FILING DATE: 08-MAR-1994  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Ferraro, Gregory D.  
: REGISTRATION NUMBER: 36,134

```

: REFERENCE/DOCKET NUMBER: 325800-465
:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201-994-1700
: TELEFAX: 201-994-1744
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1565 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 68..1207
:
: US-08-468-946-1

```

```

: RESULT 10
:
: US-08-915-096A-1
:   : Sequence 1, Application US/08915096A
:   : Patent No. 6265543
:   : GENERAL INFORMATION:
:   : APPLICANT: Weissner, Paul S.
:   : APPLICANT: Fuldner, Rebecca A.
:   : APPLICANT: Adams, Mark D.
:   : TITLE OF INVENTION: Transforming Growth Factor Alpha HI
:   : NUMBER OF SEQUENCES: 15
:   : CORRESPONDENCE ADDRESS:
:   : ADDRESSEE: Human Genome Sciences, Inc.
:   : STREET: 9410 Key West Avenue
:   : CITY: Rockville
:   : STATE: MD
:   : COUNTRY: USA
:   : ZIP: 20850
:
: COMPUTER READABLE FORM:
:   : MEDIUM TYPE: Floppy disk
:   : COMPUTER: IBM PC compatible
:   : OPERATING SYSTEM: PC-DOS/MS-DOS
:   : SOFTWARE: Patent in Release #1.0, Version #1.30
:   : CURRENT APPLICATION DATA:
:   : APPLICATION NUMBER: US/08/915,096A
:   : FILING DATE: 20-AUG-1997
:   : CLASSIFICATION: 435
:   : PRIORITY APPLICATION DATA:
:   : APPLICATION NUMBER: US 08/468,846
:   : FILING DATE: 06-JUN-1995
:   : PRIORITY APPLICATION DATA:
:   : APPLICATION NUMBER: US 08/208,008
:   : FILING DATE: 08-MAR-1994
:   : ATTORNEY/AGENT INFORMATION:
:   : NAME: Brookes, A. Anders
:   : REGISTRATION NUMBER: 36,373
:   : REFERENCE/DOCKET NUMBER: PF110D1
:   : TELECOMMUNICATION INFORMATION:
:   : TELEPHONE: 301-309-8504
:   : TELEFAX: 301-309-8439
:

```



```

Best Local Similarity 38.3%; Pred. No. 0.0043;
Matches 41; Conservative 20; Mismatches 46; Indels 0; Gaps 0;

QY 160 aatgaacttaagtgcaccaaataatagaccctgtctgtgggaactgatggaataact 219
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 1 AARAAAGTGYGCNTGYCCNAARATHYNAARCCNGTNTGYGWNWSNGAYGGNMGHACN 60

QY 220 tatcccaatgaatgcgtgtattgtttgaaggtcggaaacccagac 266
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 61 TAYGCNAAYSWNTGYATHGCGNNGTNGYAAAYGGNGTNGTNSNATHAARWS 107

RESULT 13
US-08-586-676E-6
; Sequence 6, Application US/08586676E
; Patent No. 5972698
; GENERAL INFORMATION:
; APPLICANT: Fritz, Hans,
; APPLICANT: Sommerhoff, Christian
; TITLE OF INVENTION: Trypsin Inhibitor
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5972698artis Corporation, Patent and Trademark Department
; STREET: 564 Morris Avenue
; CITY: Summit
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07901-1027
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/586,676E
; FILING DATE: 25-JAN-1996
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/02445
; FILING DATE: 25-JUL-1994
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: EP 93111930.9
; FILING DATE: 26-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pfeiffer, Hesna J.
; REGISTRATION NUMBER: 22,640
; REFERENCE/DOCKET NUMBER: 4-20076/PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 522 6940
; TELEFAX: (908) 522 6955
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Hirudo medicinalis
; US-08-586-676E-6

Query Match 9.6%; Score 35.4; DB 2; Length 138;
Best Local Similarity 38.3%; Pred. No. 0.0044;
Matches 41; Conservative 20; Mismatches 46; Indels 0; Gaps 0;

QY 160 aatgaacttaagtgcaccaaataatagaccctgtctgtgggaactgatggaataact 219
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 1 AARAAAGTGTGYCCTGYCCNAARATHYNAARCCNGTNTGYGWNWSNGATGGNMGHACN 60

QY 220 tatcccaatgaatgcgtgtattgtttgaaggtcggaaacccagac 266
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

Search completed: July 29, 2002, 17:54:42  
Job time: 5103 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 19:20:56 ; Search time 4748.22 Seconds  
(without alignments)  
1676.764 Million cell updates/sec

Title: US-09-880-107-3847  
Perfect score: 368  
Sequence: 1 gaagagacgtgtaagtcg.....aataaatgtatctgaatc 368

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues 43959072

Total number of hits satisfying chosen parameters:  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/2/pna/pna/US06\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/pna/pna/US07\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/pna/pna/US08\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/pna/pna/US09\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/pna/pna/US10\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/pna/pna/US11\_COMB.seq.\*  
7: /cgn2\_6/ptodata/2/pna/pna/US12\_COMB.seq.\*  
8: /cgn2\_6/ptodata/2/pna/pna/US13\_COMB.seq.\*  
9: /cgn2\_6/ptodata/2/pna/pna/US14\_COMB.seq.\*  
10: /cgn2\_6/ptodata/2/pna/pna/US15\_COMB.seq.\*  
11: /cgn2\_6/ptodata/2/pna/pna/US16\_COMB.seq.\*  
12: /cgn2\_6/ptodata/2/pna/pna/US17\_COMB.seq.\*  
13: /cgn2\_6/ptodata/2/pna/pna/US18\_COMB.seq.\*  
14: /cgn2\_6/ptodata/2/pna/pna/US19\_COMB.seq.\*  
15: /cgn2\_6/ptodata/2/pna/pna/US20\_COMB.seq.\*  
16: /cgn2\_6/ptodata/2/pna/pna/US21\_COMB.seq.\*  
17: /cgn2\_6/ptodata/2/pna/pna/US22\_COMB.seq.\*  
18: /cgn2\_6/ptodata/2/pna/pna/US23\_COMB.seq.\*  
19: /cgn2\_6/ptodata/2/pna/pna/US24\_COMB.seq.\*  
20: /cgn2\_6/ptodata/2/pna/pna/US25\_COMB.seq.\*  
21: /cgn2\_6/ptodata/2/pna/pna/US26\_COMB.seq.\*  
22: /cgn2\_6/ptodata/2/pna/pna/US27\_COMB.seq.\*  
23: /cgn2\_6/ptodata/2/pna/pna/US28\_COMB.seq.\*  
24: /cgn2\_6/ptodata/2/pna/pna/US29\_COMB.seq.\*  
25: /cgn2\_6/ptodata/2/pna/pna/US30\_COMB.seq.\*  
26: /cgn2\_6/ptodata/2/pna/pna/US31\_COMB.seq.\*  
27: /cgn2\_6/ptodata/2/pna/pna/US32\_COMB.seq.\*  
28: /cgn2\_6/ptodata/2/pna/pna/US33\_COMB.seq.\*  
29: /cgn2\_6/ptodata/2/pna/pna/US34\_COMB.seq.\*  
30: /cgn2\_6/ptodata/2/pna/pna/US35\_COMB.seq.\*  
31: /cgn2\_6/ptodata/2/pna/pna/US36\_COMB.seq.\*  
32: /cgn2\_6/ptodata/2/pna/pna/US37\_COMB.seq.\*  
33: /cgn2\_6/ptodata/2/pna/pna/US38\_COMB.seq.\*  
34: /cgn2\_6/ptodata/2/pna/pna/US39\_COMB.seq.\*  
35: /cgn2\_6/ptodata/2/pna/pna/US40\_COMB.seq.\*  
36: /cgn2\_6/ptodata/2/pna/pna/US41\_COMB.seq.\*  
37: /cgn2\_6/ptodata/2/pna/pna/US42\_COMB.seq.\*  
38: /cgn2\_6/ptodata/2/pna/pna/US43\_COMB.seq.\*  
39: /cgn2\_6/ptodata/2/pna/pna/US44\_COMB.seq.\*  
40: /cgn2\_6/ptodata/2/pna/pna/US45\_COMB.seq.\*  
41: /cgn2\_6/ptodata/2/pna/pna/US46\_COMB.seq.\*  
42: /cgn2\_6/ptodata/2/pna/pna/US47\_COMB.seq.\*  
43: /cgn2\_6/ptodata/2/pna/pna/US48\_COMB.seq.\*

44: /cgn2\_6/ptodata/2/pna/pna/US49\_COMB.seq.\*  
45: /cgn2\_6/ptodata/2/pna/pna/US50\_COMB.seq.\*  
46: /cgn2\_6/ptodata/2/pna/pna/US51\_COMB.seq.\*  
47: /cgn2\_6/ptodata/2/pna/pna/US52\_COMB.seq.\*  
48: /cgn2\_6/ptodata/2/pna/pna/US53\_COMB.seq.\*  
49: /cgn2\_6/ptodata/2/pna/pna/US54\_COMB.seq.\*  
50: /cgn2\_6/ptodata/2/pna/pna/US55\_COMB.seq.\*  
51: /cgn2\_6/ptodata/2/pna/pna/US56\_COMB.seq.\*  
52: /cgn2\_6/ptodata/2/pna/pna/US57\_COMB.seq.\*  
53: /cgn2\_6/ptodata/2/pna/pna/US58\_COMB.seq.\*  
54: /cgn2\_6/ptodata/2/pna/pna/US59\_COMB.seq.\*  
55: /cgn2\_6/ptodata/2/pna/pna/US60\_COMB.seq.\*  
56: /cgn2\_6/ptodata/2/pna/pna/US61\_COMB.seq.\*  
57: /cgn2\_6/ptodata/2/pna/pna/US62\_COMB.seq.\*  
58: /cgn2\_6/ptodata/2/pna/pna/US63\_COMB.seq.\*  
59: /cgn2\_6/ptodata/2/pna/pna/US64\_COMB.seq.\*  
60: /cgn2\_6/ptodata/2/pna/pna/US65\_COMB.seq.\*  
61: /cgn2\_6/ptodata/2/pna/pna/US66\_COMB.seq.\*  
62: /cgn2\_6/ptodata/2/pna/pna/US67\_COMB.seq.\*  
63: /cgn2\_6/ptodata/2/pna/pna/US68\_COMB.seq.\*  
64: /cgn2\_6/ptodata/2/pna/pna/US69\_COMB.seq.\*  
65: /cgn2\_6/ptodata/2/pna/pna/US70\_COMB.seq.\*  
66: /cgn2\_6/ptodata/2/pna/pna/US71\_COMB.seq.\*  
67: /cgn2\_6/ptodata/2/pna/pna/US72\_COMB.seq.\*  
68: /cgn2\_6/ptodata/2/pna/pna/US73\_COMB.seq.\*  
69: /cgn2\_6/ptodata/2/pna/pna/US74\_COMB.seq.\*  
70: /cgn2\_6/ptodata/2/pna/pna/US75\_COMB.seq.\*  
71: /cgn2\_6/ptodata/2/pna/pna/US76\_COMB.seq.\*  
72: /cgn2\_6/ptodata/2/pna/pna/US77\_COMB.seq.\*  
73: /cgn2\_6/ptodata/2/pna/pna/US78\_COMB.seq.\*  
74: /cgn2\_6/ptodata/2/pna/pna/US79\_COMB.seq.\*  
75: /cgn2\_6/ptodata/2/pna/pna/US80\_COMB.seq.\*

Fred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	368	100.0	368	33	US-09-880-107-3847
2	368	100.0	368	35	US-09-954-456-91
3	368	100.0	368	35	US-09-954-456-730
4	368	100.0	368	68	US-60-298-951-5
5	344.8	93.7	378	35	US-09-948-941-198
6	344.8	93.7	396	16	US-09-248-797-17742
7	344.8	93.7	396	17	US-09-346-956-7285
8	344.8	93.7	396	34	US-09-904-703-7285
9	344.8	93.7	396	34	US-09-925-564-17742
10	344.8	93.7	604	1	PCT-US00-05989-292
11	344.8	93.7	604	34	US-09-925-297-292
12	344.8	93.7	604	30	US-09-760-477-217
13	344.8	93.7	607	30	US-09-760-492-57
14	344.8	93.7	607	30	US-09-647-801-11
15	344.8	93.7	839	56	US-60-172-373-7243
16	344.8	93.7	937	25	US-09-652-127-9525
17	344.8	93.7	937	25	US-09-652-127-9397
18	343.8	93.4	459	55	US-09-652-127-7348
19	342.6	93.1	591	25	US-09-652-914-7443
20	342.6	93.1	591	25	US-09-016-434-1250
21	340.8	92.6	432	14	US-09-534-840-9544
22	337.8	91.8	478	70	US-60-363-352-2205
23	334.8	91.0	348	75	US-09-881-797-3501
24	328	89.1	1115	37	US-10-021-698-3501
25	328	89.1	1115	37	PCT-US01-08631-27303
26	324.8	88.3	438	1	PCT-US01-30732-1358
27	322.6	87.7	379	1	US-09-823-327-5264
28	321.8	87.4	421	31	US-09-30732-4270
29	316.6	86.0	362	1	PCT-US01-30732-4270
30	298	81.0	350	17	US-09-362-510-13590
31	298	81.0	350	17	US-09-362-510A-13590



32 298 81.0 350 34 US-09-904-013-13590 Sequence 13590, A  
33 286.8 77.9 250 20 US-09-534-840-9542 Sequence 9542, Ap  
34 280.6 76.3 378 31 US-09-824-130-2835 Sequence 2835, Ap  
35 272.8 74.1 276 20 US-09-534-840-9543 Sequence 9543, Ap  
36 266.2 72.3 271 17 US-09-304-517A-54874 Sequence 54874, A  
37 266.2 72.3 271 17 US-09-371-146A-54874 Sequence 54874, A  
38 266.2 72.3 271 36 US-09-985-678-54874 Sequence 54874, A  
39 265 72.0 347 1 PCT-US01-30732-1734 Sequence 1734, Ap  
40 264.6 71.9 369 31 US-09-823-327-3238 Sequence 3238, Ap  
41 264.2 71.8 300 7 US-08-303-241-1785 Sequence 1785, Ap  
42 264.2 71.8 300 20 US-09-534-840-10001 Sequence 10001, A  
43 231.6 62.9 448 30 US-09-760-477-125 Sequence 125, Ap  
44 223 60.6 510 74 US-60-350-061-74 Sequence 74, Appl  
45 222.4 60.4 549 30 US-09-770-171-1053 Sequence 1053, Ap

ALIGNMENTS

RESULT .1  
US-09-880-107-3847  
: Sequence 3847, Application US/09880107  
: GENERAL INFORMATION:  
: APPLICANT: Horne, Darci T.  
: APPLICANT: Vockley, Joseph G.  
: APPLICANT: Scherf, Uwe  
: APPLICANT: Gene Logic, Inc.  
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
: FILE REFERENCE: 44921-5028-WO  
: CURRENT APPLICATION NUMBER: US/09/880.107  
: CURRENT FILING DATE: 2001-06-14  
: PRIOR APPLICATION NUMBER: US 60/211,379  
: PRIOR FILING DATE: 2000-06-14  
: PRIOR APPLICATION NUMBER: US 60/237,054  
: PRIOR FILING DATE: 2000-10-02  
: NUMBER OF SEQ ID NOS: 3950  
: SOFTWARE: Patent In Ver. 2.1  
: SEQ ID NO 3847  
: LENGTH: 368  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
: FEATURE:  
: OTHER INFORMATION: Genbank Accession No. Y00705  
US-09-880-107-3847

Query Match 100.0%; Score 368; DB 33; Length 368;  
Best Local Similarity 100.0%; Pred. No. 5.1e-106;  
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 gaagagacgtggttaagtgcggtgcagttttcaactgacctgtgacgcagaaacttcagcc 60  
Db 1 gaagagacgtggttaagtgcggtgcagttttcaactgacctgtgacgcagaaacttcagcc 60  
QY 61 atgaaggttaacaggaactttttcttcagtgacctgtgacctgttgagtcctatctggttaac 120  
Db 61 atgaaggttaacaggaactttttcttcagtgacctgtgacctgttgagtcctatctggttaac 120  
QY 121 actggagctgactccctgggaagagagagggccaaatgttacaatgaacttaattgagtgacc 180  
Db 121 actggagctgactccctgggaagagagagggccaaatgttacaatgaacttaattgagtgacc 180  
QY 181 aagatatagacctgtctgtggagactgatgaaataactatcccaatgaatgcgtgtta 240  
Db 181 aagatatagacctgtctgtggagactgatgaaataactatcccaatgaatgcgtgtta 240  
QY 241 tgttttgaaggtcggaaacgcagacttctatctcatcaaaaaatctgggcttgcgtga 300  
Db 241 tgttttgaaggtcggaaacgcagacttctatctcatcaaaaaatctgggcttgcgtga 300  
QY 301 gaaccaaggttttgaaatcccatcaggctaccgcgagggcctattgttgataaatgtatc 360  
Db 301 gaaccaaggttttgaaatcccatcaggctaccgcgagggcctattgttgataaatgtatc 360

QY 361 tgaataac 368  
Db 361 tgaataac 368

RESULT .2  
US-09-954-456-91  
: Sequence 91, Application US/09954456  
: GENERAL INFORMATION:  
: APPLICANT: Young, Paul  
: TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C  
: TITLE OF INVENTION: Sets  
: FILE REFERENCE: 689290-76  
: CURRENT APPLICATION NUMBER: US/09/954.456  
: CURRENT FILING DATE: 2001-09-18  
: PRIOR APPLICATION NUMBER: US/60/233,617  
: PRIOR FILING DATE: 2000-09-18  
: PRIOR APPLICATION NUMBER: US/60/234,052  
: PRIOR FILING DATE: 2000-09-20  
: PRIOR APPLICATION NUMBER: US/60/234,923  
: PRIOR FILING DATE: 2000-09-25  
: PRIOR APPLICATION NUMBER: US/60/235,134  
: PRIOR FILING DATE: 2000-09-25  
: PRIOR APPLICATION NUMBER: US/60/235,637  
: PRIOR FILING DATE: 2000-09-26  
: PRIOR APPLICATION NUMBER: US/60/235,638  
: PRIOR FILING DATE: 2000-09-26  
: PRIOR APPLICATION NUMBER: US/60/235,711  
: PRIOR FILING DATE: 2000-09-27  
: PRIOR APPLICATION NUMBER: US/60/235,720  
: PRIOR FILING DATE: 2000-09-27  
: PRIOR APPLICATION NUMBER: US/60/235,840  
: PRIOR FILING DATE: 2000-09-27  
: PRIOR APPLICATION NUMBER: US/60/235,863  
: PRIOR FILING DATE: 2000-09-27  
: NUMBER OF SEQ ID NOS: 2276  
: SOFTWARE: Patent In version 3.0  
: SEQ ID NO 91  
: LENGTH: 368  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
: OTHER INFORMATION: US-09-954-456-91

Query Match 100.0%; Score 368; DB 35; Length 368;  
Best Local Similarity 100.0%; Pred. No. 5.1e-106;  
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 gaagagacgtggttaagtgcggtgcagttttcaactgacctgtgacgcagaaacttcagcc 60  
Db 1 gaagagacgtggttaagtgcggtgcagttttcaactgacctgtgacgcagaaacttcagcc 60  
QY 61 atgaaggttaacaggaactttttcttcagtgacctgtgacctgttgagtcctatctggttaac 120  
Db 61 atgaaggttaacaggaactttttcttcagtgacctgtgacctgttgagtcctatctggttaac 120  
QY 121 actggagctgactccctgggaagagagggccaaatgttacaatgaacttaattgagtgacc 180  
Db 121 actggagctgactccctgggaagagagggccaaatgttacaatgaacttaattgagtgacc 180  
QY 161 aagatatagacctgtctgtggagactgatgaaataactatcccaatgaatgcgtgtta 240  
Db 161 aagatatagacctgtctgtggagactgatgaaataactatcccaatgaatgcgtgtta 240  
QY 241 tgttttgaaggtcggaaacgcagacttctatctcatcaaaaaatctgggcttgcgtga 300  
Db 241 tgttttgaaggtcggaaacgcagacttctatctcatcaaaaaatctgggcttgcgtga 300  
QY 301 gaaccaaggttttgaaatcccatcaggctaccgcgagggcctattgttgataaatgtatc 360  
Db 301 gaaccaaggttttgaaatcccatcaggctaccgcgagggcctattgttgataaatgtatc 360

Qy 361 tgaatc 368  
Db 361 tgaatc 368

## RESULT 3

US-09-954-456-730  
; Sequence 730, Application US/09954456  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can  
; FILE OF INVENTION: Sets  
; FILE REFERENCE: 689290-76  
; CURRENT APPLICATION NUMBER: US/09/954,456  
; PRIOR FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US/60/233,617  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US/60/234,052  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,923  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,134  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,637  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,638  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,711  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,720  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,840  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,863  
; PRIOR FILING DATE: 2000-09-27  
; NUMBER OF SEQ ID NOS: 2276  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 730  
; LENGTH: 368  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-954-456-730

Query Match 100.0%; Score 368; DB 35; Length 368;  
Best Local Similarity 100.0%; Pred. No. 5.le-106;  
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gaagagacgtgtaagtgcggtgcagtttcaactgacctctggagcagagaacttcagcc 60  
Db 1 gaagagacgtgtaagtgcggtgcagtttcaactgacctctggagcagagaacttcagcc 60  
Qy 61 atgaaggttaacagcagcatcttctctcagtgcccttggtgagctctatctgtgtaac 120  
Db 61 atgaaggttaacagcagcatcttctctcagtgcccttggtgagctctatctgtgtaac 120  
Qy 121 actggaactgacctctctggtggaactgaggaacttcaactatcccaatgaatcggtta 180  
Db 121 actggaactgacctctctggtggaactgaggaacttcaactatcccaatgaatcggtta 180  
Qy 181 aagatatgacctctgctggtggaactgaggaacttcaactatcccaatgaatcggtta 240  
Db 181 aagatatgacctctgctggtggaactgaggaacttcaactatcccaatgaatcggtta 240  
Qy 241 tgtttgaaggtcggaacccagcagacttctatctcattcaaaaaatctgggcttgctga 300  
Db 241 tgtttgaaggtcggaacccagcagacttctatctcattcaaaaaatctgggcttgctga 300  
Qy 301 gaacaaaggtttgaaatcccatcagtcacccgagggcctattgttgaataaatgtatc 360  
Db 301 gaacaaaggtttgaaatcccatcagtcacccgagggcctattgttgaataaatgtatc 360  
Qy 361 tgaatc 368

Db 361 tgaatc 368

## RESULT 4

US-60-298-951-5  
; Sequence 5, Application US/60298951  
; GENERAL INFORMATION:  
; APPLICANT: Lu, Xuanyong  
; TITLE OF INVENTION: Cell Lines for Productive Infection of  
; FILE OF INVENTION: Hepatitis B Virus  
; FILE REFERENCE: LUX01.P0001  
; CURRENT APPLICATION NUMBER: US/60/298,951  
; CURRENT FILING DATE: 2001-06-18  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 368  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-60-298-951-5

Query Match 100.0%; Score 368; DB 68; Length 368;  
Best Local Similarity 100.0%; Pred. No. 5.le-106;  
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gaagagacgtgtaagtgcggtgcagtttcaactgacctctggagcagagaacttcagcc 60  
Db 1 gaagagacgtgtaagtgcggtgcagtttcaactgacctctggagcagagaacttcagcc 60  
Qy 61 atgaaggttaacagcagcatcttctctcagtgcccttggtgagctctatctgtgtaac 120  
Db 61 atgaaggttaacagcagcatcttctctcagtgcccttggtgagctctatctgtgtaac 120  
Qy 121 actggaactgacctctctggtggaactgaggaacttcaactatcccaatgaatcggtta 180  
Db 121 actggaactgacctctctggtggaactgaggaacttcaactatcccaatgaatcggtta 180  
Qy 181 aagatatgacctctgctggtggaactgaggaacttcaactatcccaatgaatcggtta 240  
Db 181 aagatatgacctctgctggtggaactgaggaacttcaactatcccaatgaatcggtta 240  
Qy 241 tgtttgaaggtcggaacccagcagacttctatctcattcaaaaaatctgggcttgctga 300  
Db 241 tgtttgaaggtcggaacccagcagacttctatctcattcaaaaaatctgggcttgctga 300  
Qy 301 gaacaaaggtttgaaatcccatcagtcacccgagggcctattgttgaataaatgtatc 360  
Db 301 gaacaaaggtttgaaatcccatcagtcacccgagggcctattgttgaataaatgtatc 360  
Qy 361 tgaatc 368

## RESULT 5

US-09-948-941-198  
; Sequence 198, Application US/09948941  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE OF INVENTION: WITH CANCER, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO000788  
; CURRENT APPLICATION NUMBER: US/09/948,941  
; CURRENT FILING DATE: 2001-09-10  
; PRIOR APPLICATION NUMBER: 60/231,328  
; PRIOR FILING DATE:  
; NUMBER OF SEQ ID NOS: 12618  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 198  
; LENGTH: 378

; TYPE: DNA  
; ORGANISM: Human  
US-09-948-941-198

Query Match 93.7%; Score 344.8; DB 35; Length 378;  
Best Local Similarity 96.8%; Pred. No. 1.2e-98;  
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

Qy 1 gaagagacgtggttaagtcggtgagctgtttcaactgacacctctggacgcagaacttcagcc 60  
Db 1 gaagagacgtggttaagtcggtgagctgtttcaactgacacctctggacgcagaacttcagcc 60  
Qy 61 atgaagtaacagcagctcttcttcagtcgcttggcctgttgctgtatctatctgtgtaac 120  
Db 61 atgaagtaacagcagctcttcttcagtcgcttggcctgttgctgtatctatctgtgtaac 120  
Qy 121 actgagctgactccctgggaagagagccaaatgttacaatgaacttaattgagtgacc 180  
Db 121 actgagctgactccctgggaagagagccaaatgttacaatgaacttaattgagtgacc 180  
Qy 181 aagatatgacacctctgtggtgagctgagaaatacttaccatgaatgcgtgtta 240  
Db 181 aagatatgacacctctgtggtgagctgagaaatacttaccatgaatgcgtgtta 240  
Qy 241 tgtttgaagtcggaacccagacacttctatctcattcaaaaatctgggaccttctga 300  
Db 241 tgtttgaagtcggaacccagacacttctatctcattcaaaaatctgggaccttctga 300  
Qy 301 gaaccaaggttttgaatcccatcagtcacccgcgagggcc-----tattgttga 350  
Db 301 gaaccaaggttttgaatcccatcagtcacccgcgagggcc-----tattgttga 350  
Qy 351 taaatgtatctgaatc 368  
Db 361 taaatgtatctgaatc 378

RESULT 6

US-09-248-797-17742

; Sequence 17742, Application US/09248797  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE REFERENCE: 20411-764  
; CURRENT FILING DATE: 1999-02-12  
; NUMBER OF SEQ ID NOS: 48909  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 17742  
; LENGTH: 396  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-248-797-17742

Query Match 93.7%; Score 344.8; DB 16; Length 396;  
Best Local Similarity 96.8%; Pred. No. 1.3e-98;  
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

Qy 1 gaagagacgtggttaagtcggtgagctgtttcaactgacacctctggacgcagaacttcagcc 60  
Db 17 gaagagacgtggttaagtcggtgagctgtttcaactgacacctctggacgcagaacttcagcc 76  
Qy 61 atgaagtaacagcagctcttcttcagtcgcttggcctgttgctgtatctatctgtgtaac 120  
Db 77 atgaagtaacagcagctcttcttcagtcgcttggcctgttgctgtatctatctgtgtaac 136  
Qy 121 actgagctgactccctgggaagagagccaaatgttacaatgaacttaattgagtgacc 180  
Db 137 actgagctgactccctgggaagagagccaaatgttacaatgaacttaattgagtgacc 196

Qy 181 aagatatgacacctctgtggtgagctgtttcaactgacaccttcccaatgaatgcgtgtta 240  
Db 197 aagatatgacacctctgtggtgagctgtttcaactgacaccttcccaatgaatgcgtgtta 256  
Qy 241 tgtttgaagtcggaacccagacacttctatctcattcaaaaatctgggaccttctga 300  
Db 257 tgtttgaagtaacgcagacacttctatctcattcaaaaatctgggaccttctga 316  
Qy 301 gaaccaaggttttgaatcccatcagtcacccgcgagggcc-----tattgttga 350  
Db 317 gaaccaaggttttgaatcccatcagtcacccgcgagggcccttattgttga 376  
Qy 351 taaatgtatctgaatc 368  
Db 377 taaatgtatctgaatc 394

RESULT 7

US-09-346-956-7285

; Sequence 7285, Application US/09346956  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE REFERENCE: 20411-758CON1  
; CURRENT FILING DATE: 1999-07-02  
; EARLIER APPLICATION NUMBER: US 09/210,298  
; NUMBER OF SEQ ID NOS: 17812  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7285  
; LENGTH: 396  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-346-956-7285

Query Match 93.7%; Score 344.8; DB 17; Length 396;

Best Local Similarity 96.8%; Pred. No. 1.3e-98;

Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

Qy 1 gaagagacgtggttaagtcggtgagctgtttcaactgacacctctggacgcagaacttcagcc 60  
Db 17 gaagagacgtggttaagtcggtgagctgtttcaactgacacctctggacgcagaacttcagcc 76  
Qy 61 atgaagtaacagcagctcttcttcagtcgcttggcctgttgctgtatctatctgtgtaac 120  
Db 77 atgaagtaacagcagctcttcttcagtcgcttggcctgttgctgtatctatctgtgtaac 136  
Qy 121 actgagctgactccctgggaagagagccaaatgttacaatgaacttaattgagtgacc 180  
Db 137 actgagctgactccctgggaagagagccaaatgttacaatgaacttaattgagtgacc 196  
Qy 181 aagatatgacacctctgtggtgagctgtttcaactgacaccttcccaatgaatgcgtgtta 240  
Db 197 aagatatgacacctctgtggtgagctgtttcaactgacaccttcccaatgaatgcgtgtta 256  
Qy 241 tgtttgaagtcggaacccagacacttctatctcattcaaaaatctgggaccttctga 300  
Db 257 tgtttgaagtaacgcagacacttctatctcattcaaaaatctgggaccttctga 316  
Qy 301 gaaccaaggttttgaatcccatcagtcacccgcgagggcc-----tattgttga 350  
Db 317 gaaccaaggttttgaatcccatcagtcacccgcgagggcccttattgttga 376  
Qy 351 taaatgtatctgaatc 368  
Db 377 taaatgtatctgaatc 394

RESULT 8

US-09-904-703-7285

Sequence 7285, Application US/09904703  
: GENERAL INFORMATION:  
: APPLICANT: Hyseq, Inc.  
: TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
: FROM VARIOUS CDNA LIBRARIES  
: FILE REFERENCE: 20411-758CON1  
: CURRENT APPLICATION NUMBER: US/09/904,703  
: CURRENT FILING DATE: 2001-07-12  
: PRIOR APPLICATION NUMBER: 09/210,298  
: PRIOR FILING DATE: 1998-12-09  
: NUMBER OF SEQ ID NOS: 17812  
: SOFTWARE: FastSeq for Windows Version 3.0  
: SEQ ID NO 7285  
: LENGTH: 396  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
US-09-904-703-7285

Query Match 93.7%: Score 344.8; DB 34; Length 396;  
Best Local Similarity 96.8%: Pred. No. 1.3e-98;  
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

QY 1 gaagagacgtgtaagtgcggtgcaagttttcaactgacctctggacgcagaaacttcagcc 60  
DB 17 gaagagacgtgtaagtgcggtgcaagttttcaactgacctctggacgcagaaacttcagcc 76  
QY 61 atgaaggttaacaggcatctttcttcagtgccctggccctgttgagtcctatctggttaac 120  
DB 77 atgaaggttaacaggcatctttcttcagtgccctggccctgttgagtcctatctggttaac 136  
QY 121 actggagctgactccctgggaagagagagccaaatattacaatgaacttaagtgaatgcacc 180  
DB 137 actggagctgactccctgggaagagagagccaaatattacaatgaacttaagtgaatgcacc 196  
QY 181 aagatatgaacctgtctgtggagctgatggaataacttaccataatgaatgcgtgtta 240  
DB 197 aagatatgaacctgtctgtggagctgatggaataacttaccataatgaatgcgtgtta 256  
QY 241 tgtttgaaggtcggaacgacgaccttctctcattcattcaaaaactctggccttgcgtga 300  
DB 257 tgtttgaaggtcggaacgacgaccttctctcattcattcaaaaactctggccttgcgtga 316  
QY 301 gaaccaaggttttgaataccatccatcaggctcaccgcgagggcc-----tattgttga 350  
DB 317 gaaccaaggttttgaataccatccatcaggctcaccgcgagggcc-----tattgttga 376  
QY 351 taaatgtatctgaatc 368  
DB 377 taaatgtatctgaatc 394

RESULT 10  
US-09-750-492-16  
: Sequence 16, Application US/09760492  
: GENERAL INFORMATION:  
: APPLICANT: Rosen et al.  
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
: FILE REFERENCE: PTZ67  
: CURRENT APPLICATION NUMBER: US/09/760,492  
: CURRENT FILING DATE: 2001-01-16  
: PRIOR application data removed - consult PALM or file wrapper  
: NUMBER OF SEQ ID NOS: 184  
: SOFTWARE: PatentIn Ver. 2.0  
: SEQ ID NO 16  
: LENGTH: 603  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
: NAME/KEY: SITE  
: LOCATION: (584)  
: OTHER INFORMATION: n equals a.t.g, or c  
US-09-760-492-16

Sequence 7285, Application US/09904703  
: GENERAL INFORMATION:  
: APPLICANT: Hyseq, Inc.  
: TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
: FROM VARIOUS CDNA LIBRARIES  
: FILE REFERENCE: 20411-758CON1  
: CURRENT APPLICATION NUMBER: US/09/904,703  
: CURRENT FILING DATE: 2001-07-12  
: PRIOR APPLICATION NUMBER: 09/210,298  
: PRIOR FILING DATE: 1998-12-09  
: NUMBER OF SEQ ID NOS: 17812  
: SOFTWARE: FastSeq for Windows Version 3.0  
: SEQ ID NO 7285  
: LENGTH: 396  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
US-09-904-703-7285

Query Match 93.7%: Score 344.8; DB 34; Length 396;  
Best Local Similarity 96.8%: Pred. No. 1.3e-98;  
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

QY 1 gaagagacgtgtaagtgcggtgcaagttttcaactgacctctggacgcagaaacttcagcc 60  
DB 17 gaagagacgtgtaagtgcggtgcaagttttcaactgacctctggacgcagaaacttcagcc 76  
QY 61 atgaaggttaacaggcatctttcttcagtgccctggccctgttgagtcctatctggttaac 120  
DB 77 atgaaggttaacaggcatctttcttcagtgccctggccctgttgagtcctatctggttaac 136  
QY 121 actggagctgactccctgggaagagagagccaaatattacaatgaacttaagtgaatgcacc 180  
DB 137 actggagctgactccctgggaagagagagccaaatattacaatgaacttaagtgaatgcacc 196  
QY 181 aagatatgaacctgtctgtggagctgatggaataacttaccataatgaatgcgtgtta 240  
DB 197 aagatatgaacctgtctgtggagctgatggaataacttaccataatgaatgcgtgtta 256  
QY 241 tgtttgaaggtcggaacgacgaccttctctcattcattcaaaaactctggccttgcgtga 300  
DB 257 tgtttgaaggtcggaacgacgaccttctctcattcattcaaaaactctggccttgcgtga 316  
QY 301 gaaccaaggttttgaataccatccatcaggctcaccgcgagggcc-----tattgttga 350  
DB 317 gaaccaaggttttgaataccatccatcaggctcaccgcgagggcc-----tattgttga 376  
QY 351 taaatgtatctgaatc 368  
DB 377 taaatgtatctgaatc 394

RESULT 9  
US-09-925-564-17742  
: Sequence 17742, Application US/09925564  
: GENERAL INFORMATION:  
: APPLICANT: Hyseq, Inc.  
: TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
: FROM VARIOUS CDNA LIBRARIES  
: FILE REFERENCE: 20411-764  
: CURRENT APPLICATION NUMBER: US/09/925,564  
: CURRENT FILING DATE: 2001-08-09  
: PRIOR APPLICATION NUMBER: 09/248,797  
: PRIOR FILING DATE: 1999-02-12  
: NUMBER OF SEQ ID NOS: 48909  
: SOFTWARE: FastSeq for Windows Version 3.0  
: SEQ ID NO 17742  
: LENGTH: 396  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
US-09-925-564-17742

|||||  
Db 357 aagatatgacccgctgctgaggactgagaaatacttatcccaatgaatgcgtgta 415  
Oy 241 tgtttgaagtcggaacccagactctctctcattcaaaatctgggcttgctga 300  
Db 417 tgtttgaagtcggaacccagactctctcattcaaaatctgggcttgctga 476  
Oy 301 gaaccaagttttgaaatcccatcaggtcaaccgcgagcc-----tattgtgaa 350  
Db 477 gaaccaagttttgaaatcccatcaggtcaaccgcgagcc-----tattgtgaa 536  
Oy 351 taaatgtatctgaatc 368  
Db 537 taaatgtatctgaatc 554

## RESULT 11

PCT-US00-05989-292  
; Sequence 292, Application PC/TUS0005989  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; APPLICANT: Steve Ruben  
; TITLE OF INVENTION: Human Pancreas and Pancreatic Cancer Associated Gene Sequences and  
; FILE REFERENCE: P105PCT  
; CURRENT APPLICATION NUMBER: PCT/US00/05989  
; EARLIER FILING DATE: 2000-03-08  
; EARLIER FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 928  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 292  
; LENGTH: 604  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: {557}  
; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: {580}  
; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: {582}  
; OTHER INFORMATION: n equals a,t,g, or c  
PCT-US00-05989-292

Query Match 93.7%; Score 344.8; DB 1; Length 604;  
Best Local Similarity 96.8%; Pred. No. 1.5e-98;  
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;  
Oy 1 gaagagacgtgtaagtcggtgcagtttcaactgacctctgacacgaacttcagcc 60  
Db 179 gaagagacgtgtaagtcggtgcagtttcaactgacctctgacacgaacttcagcc 238  
Oy 61 atgaaggttaacaggcatcttctcagtcgcttgccctgttgagctatcgtgtaac 120  
Db 239 atgaaggttaacaggcatcttctcagtcgcttgccctgttgagctatcgtgtaac 298  
Oy 121 actggagctgactccctgggaagagagccaaatgttaacaatgaacttaattggtgacc 180  
Db 299 actggagctgactccctgggaagagagccaaatgttaacaatgaacttaattggtgacc 358  
Oy 181 aagatatagacctgtctgtggagactgatgaaatactatcccaatgaatgcgtgta 240  
Db 359 aagatatagacctgtctgtggagactgatgaaatactatcccaatgaatgcgtgta 418  
Oy 241 tgtttgaaggtcggaacccagactctctcattcaaaatctgggcttgctga 300  
Db 419 tgtttgaaggtcggaacccagactctctcattcaaaatctgggcttgctga 478  
Oy 301 gaaccaagttttgaaatcccatcaggtcaaccgcgagcc-----tattgtgaa 350  
Db 479 gaaccaagttttgaaatcccatcaggtcaaccgcgagcc-----tattgtgaa 536

Oy 301 gaaccaagttttgaaatcccatcaggtcaaccgcgagcc-----tattgtgaa 350  
Db 479 gaaccaagttttgaaatcccatcaggtcaaccgcgagcc-----tattgtgaa 536  
Oy 351 taaatgtatctgaatc 368  
Db 539 taaatgtatctgaatc 556

## RESULT 12

US-09-925-297-292  
; Sequence 292, Application US/09925297  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: P105  
; CURRENT APPLICATION NUMBER: US/09/925.297  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05989  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 928  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 292  
; LENGTH: 604  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: {557}  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: {580}  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: {582}  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-297-292

Query Match 93.7%; Score 344.8; DB 34; Length 604;  
Best Local Similarity 96.8%; Pred. No. 1.5e-98;  
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

Oy 1 gaagagacgtgtaagtcggtgcagtttcaactgacctctgacacgaacttcagcc 60  
Db 179 gaagagacgtgtaagtcggtgcagtttcaactgacctctgacacgaacttcagcc 238  
Oy 61 atgaaggttaacaggcatcttctcagtcgcttgccctgttgagctatcgtgtaac 120  
Db 239 atgaaggttaacaggcatcttctcagtcgcttgccctgttgagctatcgtgtaac 298  
Oy 121 actggagctgactccctgggaagagagccaaatgttaacaatgaacttaattggtgacc 180  
Db 299 actggagctgactccctgggaagagagccaaatgttaacaatgaacttaattggtgacc 358  
Oy 181 aagatatagacctgtctgtggagactgatgaaatactatcccaatgaatgcgtgta 240  
Db 359 aagatatagacctgtctgtggagactgatgaaatactatcccaatgaatgcgtgta 418  
Oy 241 tgtttgaaggtcggaacccagactctctcattcaaaatctgggcttgctga 300  
Db 419 tgtttgaaggtcggaacccagactctctcattcaaaatctgggcttgctga 478  
Oy 301 gaaccaagttttgaaatcccatcaggtcaaccgcgagcc-----tattgtgaa 350  
Db 479 gaaccaagttttgaaatcccatcaggtcaaccgcgagcc-----tattgtgaa 536  
Oy 351 taaatgtatctgaatc 368  
Db 539 taaatgtatctgaatc 556

**RESULT 13**

```

US-09-760-477-217/C
;
; Sequence 217, Application US/09760477
;
; GENERAL INFORMATION:
;
; APPLICANT: Rosen et al.
;
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;
; FILE REFERENCE: PC013
;
; CURRENT APPLICATION NUMBER: US/09/760,477
;
; CURRENT FILING DATE: 2001-01-16
;
; Prior application data removed - consult PALM or file wrapper
;
; NUMBER OF SEQ ID NOS: 734
;
; SOFTWARE: PatentIn Ver. 2.0
;

```

Query Match 93.7%; Score 344.8; DB 30; Length 607;  
Best Local Similarity 96.8%; Pred. No. 1.5e-98;  
Matches 366: Conservative 0; Mismatches 2; Indels 10;

Qy	1	gaagagacgtggtgaatgcggtgcagttttcaactgacctctggacgcagaacttcagcc	50
Db	429	GAAGAGACGTGGTAAGTGCGGTGCAGTTTCAACTGACCTCTGGACGCAGAACTTCAGCC	370
Qy	51	atgaaggtaacaggcatctttcttcagtgacctggcccttgagctctatctcgttaac	120
Db	369	ATGAAGGTAAACAGGCATCTTCTTCAGTGGCTTGGCCCTGTGAGTCTATCTGGTAAC	310
Qy	121	actgagctgaactccctgggaagagagggccaaatgttacaatgaacttaatggatgcacc	180
Db	309	ACTGGAGCTGACTCCCTGGGACAGAGGCCAAATGTTACAATGAACTTAATGGATGCACC	250
Qy	181	aagatatatgacctgctctgctgggactgatgaaatacttatcccaatgaatgcgdtgta	240
Db	249	AAGATATATGACCCGTGCTCTGGGACTGATGAAATACCTATATCCCAATGAATGCCGTGTA	190
Qy	241	tgttttgaaggtcggaaacgcagaccttctatctctattcaaaaactctgggccttcgta	300
Db	189	TGTTTTGAAAATCGGAAACGCCAGACTTCTATCCTCATTCAAAAATCTGGGCCTTCTGTA	130
Qy	301	gaaccaaggttttgaatcccatcaggtcacgcgcgagcc-----tatgtgtgaa	350
Db	129	GAACCAAGGTTTGAAATCCCATCAGGTACACCGGAGGCCCTGACTGGCCCTTATTGTIGAA	70
Qy	351	taaatgtatctgaatatc	368
Db	69	TAAATGTATCTGAATATC	52

## RESULT 14

RESULT 14  
 US-09-760-492-57/c  
 ; Sequence 57, Application US/09760492  
 ; GENERAL INFORMATION:  
 ; TITLE OF INVENTION: all  
 ; FILE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: pr267  
 ; CURRENT APPLICATION NUMBER: US/09760,492  
 ; CURRENT FILING DATE: 2001-01-16

```

; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: PatentIn ver. 2.0

```

Query Match	93.7%	Score 344.8;	DB 30;	Length 607;
Best Local Similarity	96.8%	Pred. No. 1.5e-98;		
Matches 366. Conservative	0;	Mismatches 2;	Indels 10;	Gaps 1;

Qy	1	gaagagacgtgtaagtgcggtgcagtttccaactgaacctctggacgcgagaacttcagcc	60
Db	429	GAACGAGACGIGTAAGTGCGGTGCGAGTGTTCACATGCACTCTGGACGCAGAACTTCAGCC	370
Qy	61	atgaaggtaacagcgcatctttcttcagtcgcttggccctgttgagtcctatctcgtgaac	120
Db	369	ATGAAGGTAAACAGGCATCTTCTTCAGTGCCTTGCCCTGTGTGAGTCATCTGGTAAC	310
Qy	121	actgagcgtgactccctgggaagagagccaaatgttcacatgaacttaatggatgcacc	180
Db	309	ACTGGAGCTGACTCCCTGGGAAGAGAGGCCAAATGTTACAATGAACCTTAATGGATGCACC	250
Qy	181	aagatatatgaccctgtctgtggagctgatggaatacttatccccaatgaatcgcgtgta	240
Db	243	AAGATATATGACCCTGTCTGTGGGACTGATGGAATACTTATCCCAATGAATCGGTGTTA	190
Qy	241	tgtttgaaggtcggaaacgcgcagacttctatctcaattcaaaaactcgggccttcgtga	300
Db	189	TGTTTGTGAAAATCGGAACGGCAGCACTTCATCTCAATCAAAAATCTGGGCCCTTCCTGA	130
Qy	301	gaaccaaggtttgaaatcccatcaggtcacgcgcgagccc-----tatgttgtaa	350
Db	129	GAACCAAGGTTTGNAAATCCCATCAGGTCAACCGGAGGCCCTGACTGGCCTTATTGTTGAA	70
Qy	351	taaatgtatctgaatc	368
Db	69	TAAATGTATCTGAATATC	52

RESULT 15

```

RES001 13
US-09-647-801-11
: Sequence 11, Application US/09647801
: GENERAL INFORMATION:
: APPLICANT: Specht, Thomas
: TITLE OF INVENTION: Human Nucleic Acid Sequences which are Overexpressed
: FILE REFERENCE: 51581AWOM1XX24-P
: CURRENT APPLICATION NUMBER: US/09/647,801
: CURRENT FILING DATE: 2000-10-05
: NUMBER OF SEQ ID NOS: 71
: SEQ ID NO 11
: LENGTH: 608
: TYPE: DNA
: ORGANISM: homo sapiens
US-09-647-801-11

```

Query Match	93.7%	Score 344.8;	DB 25;	Length 608;
Best local Similarity	96.8%	Pred. No. 1.5e-98;		

us-09-880-107-3847.rnps

Tue Jul 30 07:47:49 2002

	Matches	366;	Conservative	0;	Mismatches	2;	Indels	10;	Gaps	1;
QY	1	gaagagacgtggtgaagtcggtgagtttcaactgacctctggagcgagaaacttcagcc	60							
Db	204	gaagagacgtggtgaagtcggtgagtttcaactgacctctggagcgagaaacttcagcc	263							
QY	61	atgaaggtaaacagcagcatcttctctcagtcgcttggccctgttgagctctatctggttaac	120							
Db	264	atgaaggtaaacagcagcatcttctctcagtcgcttggccctgttgagctctatctggttaac	323							
QY	121	actggagctgactccctggggaagagagcccaaatgttacaaatgaacttaaatgaatgcacc	180							
Db	324	actggagctgactccctggggaagagagcccaaatgttacaaatgaacttaaatgaatgcacc	383							
QY	181	aagatatagaccctgtctgtgggactgatggaatacttatcccaatgaatgcgtgta	240							
Db	384	aagatatagaccctgtctgtgggactgatggaatacttatcccaatgaatgcgtgta	443							
QY	241	tgtttgaagggtcggaacgccagactctctatccctcattcaaaaatctgggcttgcctga	300							
Db	444	tgtttgaagggtcggaacgccagactctctatccctcattcaaaaatctgggcttgcctga	503							
QY	301	gaaccaaggtttgaaatcccatcaggtcacccgagagcc-----tatgtgtgaa	350							
Db	504	gaaccaaggtttgaaatcccatcaggtcacccgagagcc-----tatgtgtgaa	563							
QY	351	taaatgtatctgaatc	368							
Db	564	taaatgtatctgaatc	581							

Search completed: July 29, 2002, 19:20:58  
Job time: 9799 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 18:01:11 : Search time 380.93 Seconds  
(without alignments)  
1658.637 Million cell updates/sec

Title: US-09-880-107-3847  
Perfect score: 368  
Sequence: 1 gaagagcgtgtgaagtgcg.....aataaagtatctgaatc 368

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues 3472872  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_032802.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	344.8	93.7	604	21 AAC99064	Human pancreatic c
2	344.8	93.7	608	20 AAZ41396	Human normal pancr
3	325.2	88.4	573	21 AAC04140	Human secreted pro
4	325.2	88.4	621	21 AAC04139	Human secreted pro
5	324.8	88.3	438	23 AAS91499	DNA encoding novel
6	307.8	83.6	341	9 AAN80460	Human pancreatic s
7	211.4	57.4	265	21 AAA77776	cDNA encoding huma
8	211.4	57.4	265	22 AA128514	Colon tumour relat
9	191.8	52.1	324	23 AAS91498	DNA encoding novel

10	173.8	47.2	238	18 AAT84523	hpSTI.M2 gene for
11	168.8	45.9	179	9 AAN81535	Human pancreatic s
12	166.2	45.2	171	11 AAQ01494	Modified human pan
13	166.2	45.2	171	11 AAQ01577	Modified human pan
14	162.2	44.1	420	21 AAA77777	cDNA encoding huma
15	162.2	44.1	420	21 AAA77777	Colon tumour relat
16	162.2	44.1	420	22 AA128515	HindIII-BamHI frag
17	160.8	43.7	302	15 AAQ78611	Sequence encoding
18	160.8	43.7	234	12 AAQ14262	Pancreas secretory
19	159.2	43.3	200	10 AAN91165	Lambda phage cll s
20	159.2	43.3	200	12 AAQ14261	Pancreas secretory
21	153.8	41.8	253	18 AAT84525	hpSTI.OMTKY3 chime
22	152.2	41.4	589	18 AAT84524	hpSTI.SSI chimeric
23	149	40.5	253	18 AAT84526	hpSTI.SSImutant ch
24	128	34.8	287	21 AAA87376	Rat hepatocyte car
25	109	29.6	415	17 AAT15983	PSKAN8 fragment #1
26	104.2	28.3	482	13 AAQ24083	Secretion plasmid
27	103.8	28.2	182	9 AAN80030	Synthetic DNA enco
28	103.8	28.2	182	9 AAN82217	Synthetic DNA enco
29	103.8	28.2	182	9 AAN82230	Synthetic DNA enco
30	102.2	27.8	182	9 AAN82229	Synthetic DNA enco
31	102.2	27.8	314	12 AAQ12109	Sequence encoding
32	100.8	27.4	168	9 AAN80453	Synthetic DNA enco
33	100.6	27.3	182	9 AAN82216	Synthetic DNA enco
34	100.6	27.3	182	9 AAN82228	Synthetic DNA enco
35	100.6	27.3	182	9 AAN82232	Synthetic DNA enco
36	99	26.9	182	9 AAN82221	Synthetic DNA enco
37	99	26.9	182	9 AAN82227	Synthetic DNA enco
38	99	26.9	182	9 AAN82223	Synthetic DNA enco
39	99	26.9	182	9 AAN82224	Synthetic DNA enco
40	97.4	26.5	182	9 AAN82225	Synthetic DNA enco
41	97.4	26.5	182	9 AAN82226	Synthetic DNA enco
42	97.4	26.5	182	9 AAN82227	Synthetic DNA enco
43	97.4	26.5	182	9 AAN82228	Synthetic DNA enco
44	97.4	26.5	182	9 AAN82234	Synthetic DNA enco
45	96.4	26.2	289	10 AAN90717	DNA of c1131-PSII-

ALIGNMENTS

RESULT 1  
AAC99064  
ID AAC99064 standard; cDNA; 604 BP.  
XX AC AAC99064;  
XX AC  
XX 09-MAR-2001 (first entry)  
XX Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:292.  
XX Human: pancreas; pancreatic cancer; pancreatic cancer antigen;  
XX detection; diagnosis; identification; cytostatic; neuroprotective;  
XX noctropic; immunomodulatory; relaxant; contraceptive; gynaecological;  
XX antitumour; cardiast; gene therapy; chromosome mapping;  
XX linkage analysis; tissue identification; tissue typing; forensic;  
XX neural; immune system; muscular; reproductive; gastrointestinal;  
XX pulmonary; cardiovascular; renal; proliferative; ss.  
XX Homo sapiens.  
XX WO200055320-A1.  
XX 21-SEP-2000.  
XX 08-MAR-2000; 2000WO-US05989.  
XX 12-MAR-1999; 99US-0124270.  
XX (H:MA-) HUMAN GENOME SCI INC.  
XX Rcsen CA, Ruben SM;  
XX



DR MPI; 2000-579444/54.  
XX P-PSDB; AAB54299.  
XX  
XX New nucleic acid that is a pancreatic cancer antigen for preventing,  
PT treating, or ameliorating a medical condition, particular pancreatic  
PT cancer, or for use in assays for diagnosing a pathological condition -  
XX  
XX  
XX Claim 1: Page 723-724; 1379pp; English.  
XX  
XX AAC98773 to AAC99231 encode the human pancreatic cancer associated  
CC proteins, called pancreatic cancer antigens, given in AAB54208 to  
CC AAB54466. The human pancreatic cancer antigens have cytostatic,  
CC neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,  
CC gynaecological, cardiant and antiinflammatory activities, and can be used  
CC in gene therapy. The polynucleotide and proteins can be used for  
CC preventing, treating, or ameliorating a medical condition or in assays  
CC for diagnosing a pathological condition or a susceptibility to one in a  
CC subject. Binding partners to the proteins and the activity of the  
CC proteins can be identified. The pancreatic cancer antigens can be used to  
CC detect, treat or prevent pancreatic disorders, especially cancer.  
CC Agonists and antagonists to the antigens can be screened for. The  
CC pancreatic cancer antigen polynucleotides can be used to design nucleic  
CC acid hybridisation probes that can be used in chromosome mapping, linkage  
CC analysis, tissue identification and/or typing and a variety of forensic  
CC and diagnostic methods. The proteins can be used to generate antibodies  
CC which are used to purify, detect and target the polypeptides, including  
CC both in vivo and in vitro diagnostic and therapeutic methods. The  
CC proteins can be used to treat or prevent neural, immune system, muscular,  
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or  
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent  
CC sequences used in the exemplification of the present invention.  
XX  
XX Sequence 604 BP: 169 A; 134 C; 143 G; 154 T; 4 other;

Query Match 93.7%; Score 344.8; DB 21; Length 604;  
Best Local Similarity 96.8%; Pred. No. 8.1e-108;  
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

OY 1 gaagagacgtgtaagtcggtgcagtttcaactgacctgtgacgcagacgaacttcagcc 60  
DB 179 gaagagacgtgtaagtcggtgcagtttcaactgacctgtgacgcagacgaacttcagcc 238  
OY 61 atgaaggtaacagcagcatctttcttcagtcgacctggccctgttgagtcctatctgtgtaac 120  
DB 239 atgaaggtaacagcagcatctttcttcagtcgacctggccctgttgagtcctatctgtgtaac 298  
OY 121 actggagctgactccctgggaagagagagccaaatgtttacaaatgaatgaatgcacc 180  
DB 299 actggagctgactccctgggaagagagagccaaatgtttacaaatgaatgaatgcacc 358  
OY 181 aagatatatgacctgtctgtggagctgatggaaataacttatcccaatgaatgcgtgta 240  
DB 359 aagatatatgacctgtctgtggagctgatggaaataacttatcccaatgaatgcgtgta 418  
OY 241 tgttttgaagtcggaacgcagacacttctctcattccaaatactgggaccttgcgtga 300  
DB 419 tgttttgaagtcggaacgcagacacttctctcattccaaatactgggaccttgcgtga 478  
OY 301 gaacaaagttttgaaatcccatcaggtcacgcgagggc-----tattgttgaa 350  
DB 479 gaacaaagttttgaaatcccatcaggtcacgcgagggcctgactggccttattgttgaa 538  
OY 351 taatgtatcgaatc 368  
DB 539 taatgtatcgaatc 556

RESULT 2  
AA241396  
ID AA241396 standard; cDNA; 608 BP.  
XX  
AC AA241396;

XX 19-JAN-2000 (first entry)  
XX Human normal pancreas tissue derived cDNA 11.  
XX  
XX Human; pancreas; cancer; treatment; anticancer; cytostatic;  
KW gene therapy; EST; expressed sequence tag; ss.  
XX  
XX Homo sapiens.  
OS  
XX DE19818598-A1.  
PN  
XX 21-OCT-1999.  
PD  
XX 19-APR-1998: 98DE-1018598.  
PF  
XX 19-APR-1998: 98DE-1018598.  
PR  
XX (META-) METAGEN GES GENOMFORSCHUNG MBH.  
PA  
XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;  
PI WPI; 1999-592039/51.  
XX  
XX New nucleic acid sequences expressed in normal pancreatic tissues, and  
DR derived polypeptides, for treatment of pancreatic cancer and  
PT identification of therapeutic agents -  
XX  
XX Identification of therapeutic agents -  
XX  
XX Claim 3: Page 53; 92pp; German.  
XX  
XX This invention describes novel cDNA sequences (A) that are highly  
CC expressed in normal human pancreatic tissue and which have anticancer  
CC and cytostatic activity. (A) are used (i) for recombinant expression of  
CC polypeptides (B) and (ii) to isolate complete genes. (B) are used (i) to  
CC identify agents suitable for treatment of pancreatic cancer; (ii)  
CC directly for treating this form of cancer (including expression from  
CC gene therapy vectors) and (iii) for generation of specific antibodies.  
CC (A) are identified by assembling ESTs (expressed sequence tags) from a  
CC particular tissue type before comparison of expression patterns. This  
CC allows a significantly longer fragment of the gene to be revealed, so  
CC from different libraries may represent different parts of the same  
CC unknown gene, distorting the estimated frequency of occurrence in a  
CC particular tissue. AA241396-241423 represent human normal pancreatic  
CC tissue derived cDNA fragments which encode the protein fragments  
CC represented in AA259893-Y59920.  
XX  
XX Sequence 608 BP: 182 A; 128 C; 148 G; 150 T; 0 other;

Query Match 93.7%; Score 344.8; DB 20; Length 608;  
Best Local Similarity 96.8%; Pred. No. 8.1e-108;  
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

OY 1 gaagagacgtgtaagtcggtgcagtttcaactgacctgtgacgcagacgaacttcagcc 60  
DB 204 gaagagacgtgtaagtcggtgcagtttcaactgacctgtgacgcagacgaacttcagcc 263  
OY 61 atgaaggtaacagcagcatctttcttcagtcgacctggccctgttgagtcctatctgtgtaac 120  
DB 264 atgaaggtaacagcagcatctttcttcagtcgacctggccctgttgagtcctatctgtgtaac 323  
OY 121 actggagctgactccctgggaagagagagccaaatgtttacaaatgaatgaatgcacc 180  
DB 324 actggagctgactccctgggaagagagagccaaatgtttacaaatgaatgaatgcacc 383  
OY 181 aagatatatgacctgtctgtggagctgatggaaataacttatcccaatgaatgcgtgta 240  
DB 384 aagatatatgacctgtctgtggagctgatggaaataacttatcccaatgaatgcgtgta 443  
OY 241 tgttttgaagtcggaacgcagacacttctctcattccaaatactgggaccttgcgtga 300  
DB 444 tgttttgaagtcggaacgcagacacttctctcattccaaatactgggaccttgcgtga 503

QY 301 gaaccaaggttttgaatcccatcaggctaccgcgagggc-----tatgttgaa 350  
|||||  
Db 504 gaaccaaggttttgaatcccatcaggctaccgcgagggcctgactgaccttatgttgaa 563  
|||||  
QY 351 taaatgtatctgaatac 368  
|||||  
Db 564 taaatgtatctgaatac 581  
|||||

RESULT 3  
AAC04140  
ID AAC04140 standard; cDNA: 573 BP.  
XX  
AC AAC04140;  
XX  
DT 06-OCT-2000 (first entry)  
XX  
DE Human secreted protein 5' EST, SEQ ID NO: 8215.  
XX  
KW Human; 5' EST: expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.  
XX  
OS Homo sapiens.  
XX  
PN EP1033401-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 21-FEB-2000; 2000EP-0200610.  
XX  
PR 26-FEB-1999; 99US-0122487.  
XX  
PA (GEST ) GENSET.  
XX  
PI Dumas Milne Edwards J, Duclert A, Giordano J;  
XX  
DR WPI; 2000-500381/45.  
XX  
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
XX  
PS Claim 1; SEQ ID 8215; 71pp + CD-ROM; English.  
XX  
CC The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
CC identified within the present sequence. The 5' ESTs were prepared from  
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
CC derived from the 5' ends of mRNAs and even in those cases where longer  
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.  
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC They are used to obtain upstream regulatory sequences and to design  
CC expression and secretion vectors.  
XX  
SQ Sequence 573 BP; 145 A; 139 C; 132 G; 150 T; 7 other;

Query Match 88.4%; Score 325.2; DB 21; Length 573;  
Best Local Similarity 93.2%; Pred. No. 4.1e-101;  
Matches 344; Conservative 4; Mismatches 20; Indels 1; Gaps 1;

QY 1 gaagagacgtgaatgcgagtgagtttcaactgacctctggacgcagaacttcagcc 60  
|||||  
Db 190 gaagagacgtgaatgcgagtgagtttcaactgacctctggacgcagaacttcagcc 249  
|||||  
QY 61 atgaaggttaacaggcatctttcttcagtgcccttgagcctgtgagctatctggtaac 120  
|||||

Db 250 atgaaggttaacaggcatctttcttcagtgcccttgagcctgtgagctatctggtaac 309  
QY 121 actgagagtcgactccctctgggaagagagccaaatgttacaatgaactaatgatgcacc 180  
|||||  
Db 310 actgagagtcgactccctctgggaagagagccaaatgttacaatgaactaatgatgcacc 369  
|||||  
QY 181 aagatatagaccctgtctgtgaggactatgaaatactatcccaatgaatgcgtgta 240  
|||||  
Db 370 aagatatagaccctgtctgtgaggactatgaaatactatcccaatgaatgcgtgta 429  
|||||  
QY 241 tgttttgaaggtcggaacgcagactctctatccctcattcaaaaatctgggcccctgtgta 300  
|||||  
Db 430 tgttttgaaggtcggaacgcagactctctatccctcattcaaaaatctgggcccctgtgta 489  
|||||  
QY 301 -gaaccaaggttttgaatcccatcaggctaccgcgagggcctattgttgaataaatgtat 359  
|||||  
Db 450 ggaacnaaggttttgaatcccatcaggctaccgcgagggcctgactgkctwatgtwga 549  
|||||  
QY 360 ctgaatac 368  
|||||  
Db 550 ataaatgtc 558  
|||||

RESULT 4  
AAC04139  
ID AAC04139 standard; cDNA: 621 BP.  
XX  
AC AAC04139;  
XX  
DT 06-OCT-2000 (first entry)  
XX  
DE Human secreted protein 5' EST, SEQ ID NO: 8214.  
XX  
KW Human; 5' EST: expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.  
XX  
OS Homo sapiens.  
XX  
PN EP1033401-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 21-FEB-2000; 2000EP-0200610.  
XX  
PR 26-FEB-1999; 99US-0122487.  
XX  
PA (GEST ) GENSET.  
XX  
PI Dumas Milne Edwards J, Duclert A, Giordano J;  
XX  
DR WPI; 2000-500381/45.  
XX  
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
XX  
PS Claim 1; SEQ ID 8214; 71pp + CD-ROM; English.  
XX  
CC The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
CC identified within the present sequence. The 5' ESTs were prepared from  
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
CC derived from the 5' ends of mRNAs and even in those cases where longer  
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.  
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC They are used to obtain upstream regulatory sequences and to design  
CC expression and secretion vectors.  
XX

SQ Sequence 621 BP; 154 A; 144 C; 144 G; 172 T; 7 other;

Query Match 88.4%; Score 325.2; DB 21; Length 621;  
Best Local Similarity 93.2%; Pred. No. 4.3e-101;  
Matches 344; Conservative 4; Mismatches 20; Indels 1; Gaps 1;

QY 1 gaagagacgtgtaagtgcgtgcagttttcaactgacctctgacgcagaaacttcagcc 60  
DB 238 gaagagacgtgtaagtgcgtgcagttttcaactgacctctgacgcagaaacttcagcc 297  
QY 61 atgaagtaacaggcattcttctcagtcgcttgccctgttgagtcctatctggttaac 120  
DB 298 atgaagtaacaggcattcttctcagtcgcttgccctgttgagtcctatctggttaac 357  
QY 121 actggagctgacctcctggagagagagccaaatgttaacatgaacttaagtgaatgcacc 180  
DB 358 actggagctgacctcctggagagagagccaaatgttaacatgaacttaagtgaatgcacc 417  
QY 181 aagatatgacctgtctgtggtgactgatggaataactttatcccaatgaatgcgtgta 240  
DB 418 aagatatgacctgtctgtggtgactgatggaataactttatcccaatgaatgcgtgta 477  
QY 241 tgtttgaaagtcgaaacgcagactctctcctcattcctcaaaaaatctggccctgctga 300  
DB 478 tgtttgaaataatcgaaacgcagactctctcctcattcctcaaaaaatctggccctgctga 537  
QY 301 -gaaccaaggttttgaatcccatcagtcacgcgagccctattgttgaataaatgtat 359  
DB 538 ggaacnaaggttttgaatcccatcagtcacgcgagccctattgttgaataaatgtat 597  
QY 360 ctgaatc 368  
DB 598 ataaatgctc 606

RESULT 5  
ID AAS91499 standard; cDNA; 438 BP.  
AC AAS91499;  
XX 13-FEB-2002 (first entry)  
XX DNA encoding novel human diagnostic protein #27303.  
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX Homo sapiens.  
XX WO200175067-A2.  
XX 11-OCT-2001.  
XX 30-MAR-2001; 2001WO-US08631.  
XX 31-MAR-2000; 2000US-0540217.  
XX 23-AUG-2000; 2000US-0649167.  
XX (HYSE-) HYSEQ INC.  
XX Drmanac RT, Liu C, Tang YT;  
XX WPI; 2001-639362/73.  
XX P-PSDB; ABG27312.  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity  
XX Claim 1; SEQ ID NO 27303; 103pp; English.  
PS

XX

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.  
Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 438 BP; 115 A; 100 C; 105 G; 118 T; 0 other;

Query Match 88.3%; Score 324.8; DB 23; Length 438;  
Best Local Similarity 95.0%; Pred. No. 4.9e-101;  
Matches 361; Conservative 0; Mismatches 7; Indels 12; Gaps 2;

QY 1 gaagagacgtgtaagtgcgtgcagttttcaactgacctctggacgcagaaacttcagcc 60  
DB 52 gaagagacgtgtaagtgcgtgcagttttcaactgacctctggacgcagaaacttcagcc 111  
QY 61 atgaagtaacaggcattcttctcagtcgcttgccctgttgagtcctatctggttaac 120  
DB 112 atgaagtaacaggcattcttctcagtcgcttgccctgttgagtcctatctggttaac 171  
QY 121 actggagctgacctcctggagagagagccaaatgttacaatgaacttaagtgaatgcacc 180  
DB 172 actggagctgacctcctggagagagagccaaatgttacaatgaacttaagtgaatgcacc 231  
QY 181 aagatatgacctgtctgtggtgactgatggaataactttatcccaatgaatgcgtgta 240  
DB 232 aagatatgacctgtctgtggtgactgatggaataactttatcccaatgaatgcgtgta 291  
QY 241 tgtttgaaagtcgaaacgcagactctctcctcattcctcaaaaaatctggccctgctga 300  
DB 292 tgtttgaaataatcgaaacgcagactctctcctcattcctcaaaaaatctggccctgctga 351  
QY 301 gaacca-agggtttgaaatcccatcaggtcacgcg- - - - -aggcctattgttg 348  
DB 352 gaaccagaggtttgaaatcccatcaggtcacgcgagggcctgactgacctattgttg 411  
QY 349 aataaatgtatcgaatc 368  
DB 412 aataaatgtatcgaatc 431

RESULT 6  
ID AAN80460 standard; DNA; 341 BP.  
AC AAN80460;  
XX 23-NOV-1990 (first entry)  
XX Human pancreatic secretory trypsin inhibitor-encoding sequence.  
XX human pancreatic secretory trypsin inhibitor (PSTI);  
KW pancreatic diseases; ss.  
XX Homo sapiens.  
OS

```

XX FH Key Location/Qualifiers
XX CDS 26..265
XX FT /*tag= a
XX FT 26..94
XX FT sig_peptide
XX FT /*tag= b
XX FT mat_peptide
XX FT 95..262
XX FT /*tag= c
XX FT /product=PSTI

XX EP267692-A.
XX PN
XX XX
XX PD 18-MAY-1988.
XX PF 14-OCT-1987; 87EP-0309087.
XX PR 14-MAY-1986; 86JP-0245049.
XX PA (SHIO ) SHIONOGI SEIYAKU KK.
XX PI Ogawa M, Matsubara K;
XX DR WPI; 1988-134544/20.
XX DR P-PSDB; AAP80460.
XX PT Recombinant human pancreatic secretory trypsin inhibitor
XX PT used for in vitro diagnosis and monitoring of pancreatic
XX PT diseases.
XX PS Disclosure; ; p: English.
XX XX
XX CC Saccharomyces cerevisiae AH22 were transformed with a vector contg
XX CC this sequence. The cDNA was isolated from a pBR322 library in
XX CC E.coli K12 HB101. cDNA was prepared from mRNA obtained from human
XX CC pancreas. The library was screened firstly by a probe to codons
XX CC specifying amino acids 8-12 of mature PSTI. DNA fragments
XX CC identified in this way were then used to re-screen the library for
XX CC the full-length gene. Recombinant PSTI was obtained upon culturing
XX CC yeast transformants.
XX CC The PSTI-encoding sequence can also be chemically synthesised.
XX CC See also AAN82422.
XX SQ Sequence 341 BP; 94 A; 76 C; 78 G; 93 T; 0 other;

Query Match 83.6%; Score 307.8; DB 9; Length 341;
Best Local Similarity 96.5%; Pred. No. 2.9e-95;
Matches 329; Conservative 0; Mismatches 2; Indels 10; Gaps 1:

Oy 36 gaccttggagcgcagaaacttcagccatgaaggttaacaggcattcttcttcagtcgctt 95
Db 1 gaccttggagcgcagaaacttcagccatgaaggttaacaggcattcttcttcagtcgctt 60

Oy 96 ggccttggagtcctatctgttaacactggagtcgactccctgggaagagagccaaatg 155
Db 61 ggccttggagtcctatctgttaacactggagtcgactccctgggaagagagccaaatg 120

Oy 156 ttacaatgaacttaagtgtgacccaagatatatgacctgtctgtggagactgatggaaa 215
Db 121 ttacaatgaacttaagtgtgacccaagatatatgacctgtctgtggagactgatggaaa 180

Oy 216 tacttatcccaatgaatgcgtgtatgttttgaagtcggaacgcagacttctatcct 275
Db 181 tacttatcccaatgaatgcgtgtatgttttgaagtcggaacgcagacttctatcct 240

Oy 276 cattcaaaaatctggccttgcgtgagaaccaaagttttgaaatcccatcagggtcacccg 335
Db 241 cattcaaaaatctggccttgcgtgagaaccaaagttttgaaatcccatcagggtcacccg 300

Oy 336 aggcc-----tatttgaataaatgtatctgaata 366
Db 301 aggcctgactggccttatttgaataaatgtatctgaata 341

```

```

RESULT 7
AAA77776
ID AAA77776 standard; cDNA: 265 BP.
XX AC
XX AA77776;
XX DT 14-NOV-2000 (first entry)
XX DE cDNA encoding human colon tumour polypeptide, SEQ ID NO:55.
XX KW Human colon tumour polypeptide; tumour antigen; cancer; vaccine;
XX KW immunotherapy; diagnosis; progression; ss.
XX OS Homo sapiens.
XX PN WO200037643-A2.
XX PD 29-JUN-2000.
XX PF 23-DEC-1999; 99WO-US30909.
XX PR 23-DEC-1998; 98US-0221298.
XX PR 02-JUL-1999; 99US-0347496.
XX PR 22-SEP-1999; 99US-0401064.
XX PR 19-NOV-1999; 99US-0444242.
XX PR 02-DEC-1999; 99US-0454150.
XX PA (CORI-) CORIXA CORP.
XX PI Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk J;
XX PI Wang T, Yugu J;
XX DR WPI; 2000-442671/38.
XX PT New colon tumor polypeptides used to inhibit the development of cancer,
XX PT especially colon cancer, and for diagnosing and monitoring the
XX PT progression of the cancer.
XX PS Claim 29; Page 103; 229pp; English.
XX CC Sequences AAA77722-A78199 represent 478 cDNAs encoding proteins or
XX CC portions of proteins which are associated with human colon tumours.
XX CC The invention also specifically discloses 8 human colon tumour proteins
XX CC (AAB11897-B11904). The nucleic acids, the polypeptides they encode, and
XX CC antigen presenting cells (APCs), preferably dendritic cells) expressing
XX CC such polypeptides may be used in vaccines that target tumour cells,
XX CC especially colon tumour cells, thereby inhibiting the development of
XX CC cancer. T-cells specific for the polypeptide expressed by the APC are
XX CC used to remove tumour cells from biological samples, especially blood or
XX CC fractions thereof. The sample or the isolated T-cells specific for the
XX CC polypeptide can then be used to inhibit cancer development. CD4+ and/or
XX CC CD8+ T-cells from a patient may be incubated with a polypeptide or
XX CC nucleic acid of the invention, or an APC expressing such a polypeptide,
XX CC to cause the proliferation of specific T-cells. The T-cells can be
XX CC cloned and then administered back to the patient to inhibit cancer
XX CC development. Nucleic acids encoding the polypeptides and antibodies
XX CC against the polypeptides may be used to determine the expression level
XX CC of a tumour protein of the invention, and therefore to determine whether
XX CC cancer cells are present. Such diagnostic methods may also be used to
XX CC monitor the progression of a cancer by repeating the processes at time
XX CC intervals, and comparing the current result to previous results. The
XX CC present sequence represents a cDNA encoding a human colon tumour
XX CC polypeptide.
XX SQ Sequence 265 BP; 69 A; 59 C; 57 G; 75 T; 5 other;

```

```

Query Match 57.4%; Score 211.4; DB 21; Length 265;
Best Local Similarity 95.6%; Pred. No. 2.9e-62;
Matches 237; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

Oy 75 atctttcttcagtcgcttggcccttggatctctgttaacactgagctgactcc 135

```

|||||  
Db 1 atcttctctcagtcgcttgccctgttgagtcctatctggttaacactgagctgactcc 60  
QY 136 ctgggaagagagcccaaatgttacaatgaacttaattgagtcacacagatatatgacct 195  
Db 61 ctgggaagagagcccaaatgttacaatgaacttaattgagtcacacagatatatgacct 120  
QY 196 gtcctgagagactgagaaatacttcccaatgaatg-cgtgttatg-ttttcaaggtc 253  
Db 121 gtcctgagagactgagaaatacttcccaatgaatg-cgtgttatg-ttttcaaggtc 180  
QY 254 ggaacgcagacttctatctcctcattcaaaaatctggtggtggtggtggtggtttt 313  
Db 181 ggaacgcagacttctatctcctcattcaaaaatctggtggtggtggtggtttt 240  
QY 314 gaaatccc 321  
Db 241 naaaatcc 248

RESULT 8  
AAI28514  
ID AAI28514 standard; cDNA: 265 BP.  
XX  
AC AAI28514;  
XX  
DT 12-OCT-2001 (first entry)  
XX  
DE Colon tumour related cDNA sequence 11099.  
XX  
KW Human; immunotherapy; diagnosis; colon cancer; colon tumour;  
KW Immunogenic; gene therapy; vaccine; colonic cancer; ss.  
XX  
OS Homo sapiens.  
XX  
FN WO200149716-A2.  
XX  
PD 12-JUL-2001.  
XX  
PF 29-DEC-2000; 2000WO-US35596.  
XX  
PR 30-DEC-1999; 99US-0476296.  
PR 10-JAN-2000; 2000US-0480321.  
PR 15-FEB-2000; 2000US-0504629.  
PR 06-MAR-2000; 2000US-0519444.  
PR 19-MAY-2000; 2000US-0575251.  
PR 29-JUN-2000; 2000US-0609448.  
PR 28-AUG-2000; 2000US-0649811.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stoik JA;  
PI King GE, Wang T, Jiang Y;  
XX  
WPI; 2001-441847/47.  
XX  
PT Colon tumor associated proteins and nucleic acids useful for the  
PT prevention, diagnosis and treatment of colonic cancer -  
XX  
PS Claim 25; Page 145; 472pp; English.

XX  
CC The present invention describes colon tumour associated proteins (I) and  
CC the polynucleotides (II) that encode them. (I) have cytostatic activity.  
CC (I) and (II) can be used in gene therapy and vaccine production. (I) and  
CC (II) may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate colon tumour associated protein (TCAP)  
CC expression, such as colonic cancer. For example, (I) and (II) may be  
CC used to treat disorders associated with decreased expression by  
CC rectifying mutations or deletions in a patient's genome that affect the  
CC activity of TCAPs by expressing inactive proteins or to supplement the  
CC patients own production of them. Additionally, (II) may be used to  
CC produce the TCAP proteins, by inserting the nucleic acids into a host  
CC cell culturing the cell to express the protein. (II) and its

CC complementary sequences may also be used as DNA probes in diagnostic  
CC polymerase chain reaction (PCR) and hybridisation assays to detect and  
CC quantitate the presence of similar nucleic acids in samples, and  
CC therefore which patients may be in need of restorative therapy. (I) may  
CC also be used as antigens in the production of antibodies against TCAPs  
CC and in assays to identify modulators of TCAP expression and activity.  
CC Anti-(I) antibodies and antagonists may also be used to down regulate  
CC TCAP expression and activity. The anti-(I) antibodies may also be used  
CC as diagnostic agents for detecting the presence of TCAPs in samples  
CC (e.g. by enzyme linked immunosorbent assay (ELISA)). AAI28460 to AAI29512  
CC and AAI24494 to AAI24523 represent nucleotide and amino acid sequences  
CC given in the exemplification of the present invention.  
XX  
SQ Sequence 265 BP: 69 A; 59 C; 57 G; 75 T; 5 other;

Query Match 57.4%; Score 211.4; DB 22; Length 265;  
Best Local Similarity 95.6%; Pred. No. 2.9e-62;  
Matches 237; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 76 atcttctctcagtcgcttgccctgttgagtcctatctggttaacactgagctgactcc 135  
Db 1 atcttctctcagtcgcttgccctgttgagtcctatctggttaacactgagctgactcc 60  
QY 136 ctgggaagagagcccaaatgttacaatgaacttaattgagtcacacagatatatgacct 195  
Db 61 ctgggaagagagcccaaatgttacaatgaacttaattgagtcacacagatatatgacct 120  
QY 136 gtcctgagagactgagaaatacttcccaatgaatg-cgtgttatg-ttttcaaggtc 253  
Db 121 gtcctgagagactgagaaatacttcccaatgaatg-cgtgttatg-ttttcaaggtc 180  
QY 254 ggaacgcagacttctatctcctcattcaaaaatctggtggtggtggtggtttt 313  
Db 181 ggaacgcagacttctatctcctcattcaaaaatctggtggtggtggtttt 240  
QY 314 gaaatccc 321  
Db 241 naaaatcc 248

RESULT 9  
AAS91498  
ID AAS91498 standard; cDNA: 324 BP.  
XX  
AC AAS91498;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #27302.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US08631.  
XX  
PR 31-MAR-2000; 2000US-0540217.  
PR 25-AUG-2000; 2000US-0649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI DrTanJC RT, Liu C, Tang YT;  
XX  
WPI; 2001-639362/73.  
XX  
P-FSD3; ABG27311.  
XX  
PI New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity

PS Claim 1; SEQ ID No 27302; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 324 BP; 76 A; 68 C; 87 G; 93 T; 0 other;

Query Match 52.1%; Score 191.8; DB 23; Length 324;  
Best Local Similarity 99.0%; Pred. No. 1.6e-55;  
Matches 193; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 61 atgaaggttaacagcgtcttcttcagtcgttgccctgttgagtcctatctggtaac 120  
Db 1 atgaaggttaacagcgtcttcttcagtcgttgccctgttgagtcctatctggtaac 60  
Qy 121 actggagctgactccctggagagagagagccaaatgttacaatgaacttaagtgaacc 180  
Db 61 actggagctgactccctggagagagagagccaaatgttacaatgaacttaagtgaacc 120  
Qy 181 aagatatagacctgtctgtggagctgatgaaatactatcccaatgaatgcgtgtta 240  
Db 121 aagatatagacctgtctgtggagctgatgaaatactatcccaatgaatgcgtgtta 180  
Qy 241 tgttttgaaggtcgg 255  
Db 181 tgttttgaaggtcgg 195

RESULT 10  
AAT84523  
ID AAT84523 standard; DNA; 238 BP.

AC AAT84523;

XX 02-DEC-1997 (first entry)

XX hpSTI.M2 gene for baculoviral expression.

XX Subtilisin inhibitor; protease inhibitor; PSTI; hpSTI.M2; human;  
KW pancreatic secretory trypsin inhibitor; enzyme engineering;  
KW protein engineering; baculovirus; detergent; ss.

XX Chimeric Homo sapiens;  
OS Chimeric synthetic.

XX Key Location/Qualifiers  
FT CDS 1..228 /\*Cag- a

PN WO9715670-A1.

XX 01-MAY-1997.

XX 25-OCT-1996; 96WO-US17153.

XX 25-OCT-1995; 95US-0548186.

XX (ARRI-) ARRIS PHARM CORP.

XX Estell DA, Fisher J, Hartman C, McGrath M, Schmidt B;

PI WPI; 1997-259023/23.

XX P-PSDB; AAW26361.

XX Modified subtilisin inhibitor - having altered reactive site  
PT resulting in decreased immunogenicity, used in detergent  
PT composition(s)

PS Example 1; Fig 8; 69pp; English.

XX This DNA sequence encodes a polypeptide (AAW26361) comprising human  
CC pancreatic secretory trypsin inhibitor (PSTI), modified to include a  
CC C-terminal M2 tag. The sequence is suitable for baculovirus  
CC expression. Novel modified subtilisin inhibitors are claimed in  
CC in which a human standard mechanism inhibitor, such as PSTI, has at  
CC least one amino acid (aa) of the reactive site substituted with  
CC different aa so that its dissociation constant for inhibition (Ki),  
CC for at least one subtilisin, is reduced by at least a factor of  
CC 100. Also claimed are: a nucleic acid (I) encoding the subtilisin  
CC inhibitor; expression vectors containing (I); and host cells  
CC containing (I). The subtilisin inhibitors (see AAW26362-64) combine  
CC the low allergenicity of human standard mechanism inhibitors, which  
CC are not specific for subtilisin, and the high affinity of non-human  
CC subtilisin inhibitors. They form a complex with subtilisin, so as  
CC to avoid problems of allergenicity in detergent formulations, and  
CC can also be used in affinity purification and (diagnostic)  
CC quantification of subtilisin.

XX Sequence 238 BP; 68 A; 58 C; 61 G; 51 T; 0 other;

Query Match 47.2%; Score 173.8; DB 18; Length 238;

Best Local Similarity 96.2%; Pred. No. 2.1e-49;  
Matches 178; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 113 ctggttaacactggagctgactccctgggaagagagagccaaatgttacaatgaacttaatg 172  
Db 11 ccggcaacaactggagctgactccctgggaagagagagccaaatgttacaatgaacttaatg 70  
Qy 173 gatgaccacaagatatatgacctgtctgtggagctgatgaaatactatcccaatgaat 232  
Db 71 gatgaccacaagatatatgacctgtctgtggagctgatgaaatactatcccaatgaat 130  
Qy 233 gcgtgttatgttttgaaggctcggaacgccagactctatcctcatcctcaaaatctggcc 292  
Db 131 gcgtgttatgttttgaaggctcggaacgccagactctatcctcatcctcaaaatctggcc 190  
Qy 293 ctgtgc 297  
Db 191 ctgtgc 195

RESULT 11

AAN81535

ID AAN81535 standard; DNA; 179 BP.

XX AAN81535;

XX 07-DEC-1999 (first entry)

XX Human pancreatic secretory trypsin inhibitor synthetic gene.

human pancreatic secretory trypsin inhibitor (hPSTI); amino-glycoside 3-phosphotransferase (APT); fusion protein; ds; synthetic.

EP264118-A.

20-APR-1988.

14-OCT-1987; 87EP-0115016.

14-OCT-1986; 86JP-0245048.

23-DEC-1986; 86JP-0314603.

(SHIO ) SHIONOGI SEIYAKU KK.

Teraoka H, Shin M, Ohara O, Kiruchi N;

WPI; 1988-106896/16.

P-PSDB; AAP81182.

Prepn. of proteins by recombinant DNA method - comprises use of vector expressing proteins as fused proteins with amino-glycosidase 3-phosphotransferase and separating prods from them.

Disclosure; : p; English.

This synthetic sequence was constructed from 10 oligonucleotides for the upper strand given (and 10 for the lower strand). The upper strand overhangs the complementary strand by 2 bases; the lower strand overhangs the upper strand by 4 bases, namely CTAAG.

This synthetic gene was fused to the amino-glycosidase 3-phosphotransferase gene and the plasmid used to transform E.coli.

PSTI is then produced by culturing the transformants to produce APT. PSTI is used in the diagnosis and observation of pancreatic diseases.

Sequence 179 BP; 54 A; 37 C; 41 G; 47 T; 0 other;

Query Match 45.9%; Score 168.8; DB 9; Length 179;

Best Local Similarity 98.8%; Pred. No. 9.4e-48;

Matches 170; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 130 gactccctgggaagagagggcccaaatgttacaatgaacttaataatgaatgcaccagatatat 189

DB 8 gactccctgggaagagagggcccaaatgttacaatgaacttaataatgaatgcaccagatatat 67

QY 190 gactccgtctgtggactgaggaataactatcccaatgaatgcgtgttatgtttgaa 249

DB 68 gactccgtctgtggactgaggaataactatcccaatgaatgcgtgttatgtttgaa 127

QY 250 ggtcggaaacccagacttctatctcattcccaaaatctgggcccctgtgag 301

DB 128 aatcggaaacccagacttctatctcattcccaaaatctgggcccctgtgag 179

RESULT 12

AAQ01494

ID AAQ01494 standard; DNA; 171 BP.

XX

AC AAQ01494;

XX

DT 04-JUL-1990 (first entry)

XX

DE Modified human pancreatic secretory trypsin inhibitor.

XX

KW Human; pancreatic secretory trypsin inhibitor

XX

OS Homo sapiens.

XX

PA (SHIO ) SHIONOGI SEIYAKU KK.

EP352089-A.

24-JAN-1990.

19-JUL-1989; 89EP-0307309.

11-OCT-1988; 88JP-0181316.

19-JUL-1988; 88JP-0255580.

(SHIO ) SHIONOGI SEIYAKU KK.

Yoshida N, Kikuchi N, Teraoka H, Shin M;

WPI; 1990-024454/04

P-PSDB; AAR03728.

Human PSTI modified at arginine 42 and/or 44 by substitution with glutamine and/or serine to allow trypsin-inhibitory activity.

Disclosure; Fig 2; 25pp; English.

This modified human pancreatic secretory trypsin inhibitor differs from natural human PSTI in that bases 130-32 that usually code for Arg are replaced with Ser. The modified PSTI allows a sustained trypsin inhibition effect useful in the clinical treatment of acute pancreatitis. See also AAC01377, AAR01493, AAR03727, AAR03728 and AAR03729.

Sequence 171 BP; 53 A; 34 C; 38 G; 46 T; 0 other;

Query Match 45.2%; Score 166.2; DB 11; Length 171;

Best Local Similarity 98.2%; Pred. No. 7.2e-47;

Matches 168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 130 gactccctgggaagagagggcccaaatgttacaatgaacttaataatgaatgcaccagatatat 189

DB 1 gactccctgggaagagagggcccaaatgttacaatgaacttaataatgaatgcaccagatatat 60

QY 150 gactccgtctgtggactgaggaataactatcccaatgaatgcgtgttatgtttgaa 249

DB 61 gactccgtctgtgtggactgaggaataactatcccaatgaatgcgtgttatgtttgaa 120

QY 250 ggtcggaaacccagacttctatctcattcccaaaatctgggcccctgtgag 300

DB 121 aatcggaaacccagacttctatctcattcccaaaatctgggcccctgtgag 171

RESULT 13

AAQ01577

ID AAQ01577 standard; DNA; 171 BP.

XX

AC AAQ01577;

XX

DT 04-JUL-1990 (first entry)

XX

DE Modified human pancreatic secretory trypsin inhibitor.

XX

KW Human; pancreatic secretory trypsin inhibitor.

XX

OS Homo sapiens.

XX

PN EP352089-A.

XX

PD 24-JAN-1990.

XX

PF 19-JUL-1989; 89EP-0307309.

XX

PR 11-OCT-1988; 88JP-0181316.

XX

PR 19-JUL-1988; 88JP-0255580.

XX

XX (SHIO ) SHIONOGI SEIYAKU KK.

XX YOSHIDA N, Kikuchi N, Teraoka H, Shin M;  
XX WPI: 1990-024464/04  
XX P-PSDB: AAR03727.  
XX Human PSTII modified at arginine 42 and/or 44  
XX by substitution with glutamine and/or serine to allow  
XX trypsin-inhibitory activity.  
XX Disclosure: fig 1: 25pp; English.  
XX This modified human pancreatic secretory trypsin inhibitor  
XX differs from natural human PSTI in that bases 124-26 that  
XX usually code for Arg are replaced with Gln. The modified  
XX PSTI allows a sustained trypsin inhibition effect useful  
XX in the clinical treatment of acute pancreatitis. See also  
XX AAQ01494, AAQ01493, AAR03727, AAR03728 and AAR03729.  
XX Sequence 171 BP; 53 A; 35 C; 37 G; 46 T; 0 other;  
SQ  
Query Match 45.2%; Score 166.2; DB 11; Length 171;  
Best Local Similarity 98.2%; Pred. No. 7.2e-47;  
Matches 168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 130 gactccctgggaagagagcccaaatgttacaaatgaacttaattggaatggaacccaagatatat 189  
DB 1 gactccctgggaagagagcccaaatgttacaaatgaacttaattggaatggaacccaagatatat 60  
QY 190 gaccctgtctgtgggactgatggaataacttatcccaatgaatgcgtgttatgtttgaa 249  
DB 61 gaccctgtctgtgggactgatggaataacttatcccaatgaatgcgtgttatgtttgaa 120  
QY 250 ggtcgaaacgcagactctctatccctcattcaaaaaatctgggacctgtgta 300  
DB 121 aatcagaacgcagactctctatccctcattcaaaaaatctgggacctgtgta 171  
RESULT 14  
AAAT7777/c  
ID AAA77777 standard; cDNA: 420 BP.  
XX AAA77777;  
XX 14-NOV-2000 (first entry)  
XX cDNA encoding human colon tumour polypeptide, SEQ ID NO:56.  
XX Human colon tumour polypeptide; tumour antigen; cancer; vaccine;  
XX immunotherapy; diagnosis; progression; ss.  
XX Homo sapiens.  
XX WO200037643-A2.  
XX 29-JUN-2000.  
XX 23-DEC-1999; 99WO-US30909.  
XX 23-DEC-1998; 98US-0221298.  
XX 02-JUL-1999; 99US-0347496.  
XX 22-SEP-1999; 99US-0401064.  
XX 19-NOV-1999; 99US-0444242.  
XX 02-DEC-1999; 99US-0454150.  
XX (CORI-) CORIXA CORP.  
XX Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stoik J;  
XX Wang T, Yuqiu J;  
XX WPI: 2000-442671/38.  
XX

PT New colon-tumor polypeptides used to inhibit the development of cancer,  
PT especially colon cancer, and for diagnosing and monitoring the  
PT progression of the cancer -  
XX Claim 29: Page 103; 229pp; English.  
XX Sequences AAA77722-A78199 represent 478 cDNAs encoding proteins or  
XX portions of proteins which are associated with human colon tumours.  
XX The invention also specifically discloses 8 human colon tumour proteins  
XX (AAB1897-B11904). The nucleic acids, the polypeptides they encode, and  
XX antigen presenting cells (APCs), preferably dendritic cells) expressing  
XX such polypeptides may be used in vaccines that target tumour cells,  
XX especially colon tumour cells, thereby inhibiting the development of  
XX cancer. T-cells specific for the polypeptide expressed by the APC are  
XX used to remove tumour cells from biological samples, especially blood or  
XX fractions thereof. The sample or the isolated T-cells specific for the  
XX polypeptide can then be used to inhibit cancer development. CD4+ and/or  
XX CD8+ T-cells from a patient may be incubated with a polypeptide or  
XX nucleic acid of the invention, or an APC expressing such a polypeptide,  
XX to cause the proliferation of specific T-cells. The T-cells can be  
XX cloned and then administered back to the patient to inhibit cancer  
XX development. Nucleic acids encoding the polypeptides and antibodies  
XX against the polypeptides may be used to determine the expression level  
XX of a tumour protein of the invention, and therefore to determine whether  
XX cancer cells are present. Such diagnostic methods may also be used to  
XX monitor the progression of a cancer by repeating the processes at time  
XX intervals, and comparing the current result to previous results. The  
XX present sequence represents a cDNA encoding a human colon tumour  
XX polypeptide.  
SQ Sequence 420 BP; 103 A; 107 C; 88 G; 111 T; 11 other;

Query Match 44.1%; Score 162.2; DB 21; Length 420;  
Best Local Similarity 96.2%; Pred. No. 2.6e-45;  
Matches 175; Conservative 0; Mismatches 6; Indels 1; Gaps 1;  
QY 160 aatgaacttaattggaatggaacccaagatatatgacccctgtctgtggactgatggaataac 218  
DB 230 AANGAACTTAATGGATGACCCCAAGATATATGACCCNGTCTGTGGGACTGATGGAATAC 141  
QY 219 ttatcccaatgaatgcgtgttatgtttgaaggctcggaacccagactctatctctcat 278  
DB 140 TTATCCCAATGAATGCGTGTATGTTTGAATAATCGGAACGCCAGACTTCTATCTCAT 81  
QY 279 tcaaaaatctgggacctgtctgtagaaccacaaagtttgaatcccatcaggtcaccgcgag 338  
DB 20 TCAAAAATCTGGGCTTCTGTGAGAACCAAGGTTTGAATCCCATCATCAGGTACCGCGAGG 21  
QY 339 AC 340  
DB 20 AC 19  
RESULT 15  
AAI28515/c  
ID AAI28515 standard; cDNA: 420 BP.  
XX AAI28515;  
XX 12-OCT-2001 (first entry)  
XX Colon tumour related cDNA sequence 11186.  
XX Human: immunotherapy; diagnosis; colon cancer; colon tumour;  
XX immunogenic; gene therapy; vaccine; colonic cancer; ss.  
XX Homo sapiens.  
XX WO200149716-A2.  
XX 12-JUL-2001.  
XX





GenCore version 4.5

Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 17:20:29 ; Search time 394.41 Seconds

(without alignments)

2492.057 Million cell updates/sec

Title: US-09-880-107-2492

Perfect score: 466

Sequence: 1 atcaagaacacatagatttc.....ggtaagggtggatgtctt 466

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1333638 seqs, 1054605264 residues

Total number of hits satisfying chosen parameters: 2667276

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA, New:\*

1: /cgn2\_6/ptodata/2/pna/US06\_NEW\_COMB.seq:\*

2: /cgn2\_6/ptodata/2/pna/US07\_NEW\_COMB.seq:\*

3: /cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq:\*

4: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:\*

5: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:\*

6: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:\*

7: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:\*

8: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	397	85.2	777	1	PCT-US02-18947-1279
c 2	397	85.2	777	7	US-10-171-311-229
c 3	397	85.2	777	7	US-10-172-118-1279
c 4	343.8	73.8	477	7	US-10-102-524-299
c 5	334.4	71.8	474	7	US-10-040-862-205
c 6	334.4	71.8	474	7	US-10-040-862-646
c 7	334.4	71.8	474	7	US-10-040-862-4647
c 8	323.4	69.4	453	7	US-10-040-862-248
c 9	323.4	69.4	453	7	US-10-040-862-5273
c 10	45.8	9.8	1274	7	US-10-137-036-34
c 11	44.2	9.5	442	5	US-09-620-393B-2769
c 12	44.2	9.5	501	5	US-09-620-393B-4443
c 13	43.4	9.3	243	5	US-09-975-254-22419
c 14	42.6	9.1	552	5	US-09-620-393B-1362
c 15	42.2	9.1	469	5	US-09-620-393B-868
c 16	42.2	9.1	469	5	US-09-620-393B-18893
c 17	41	8.8	455	5	US-09-935-625-18645
c 18	41	8.8	595	5	US-09-620-393B-1166
c 19	41	8.8	596	5	US-09-620-393B-6996
c 20	39.8	8.5	554	5	US-09-886-492-17764
c 21	39.4	8.5	452	5	US-09-935-625-21348
c 22	39.4	8.5	587	5	US-09-620-393B-6747
c 23	39	8.4	461	5	US-09-620-393B-5212
c 24	39	8.4	461	5	US-09-935-625-2477
c 25	39	8.4	461	5	US-09-935-625-6719
c 26	39	8.4	461	5	US-09-935-625-11030
c 27	39	8.4	461	5	US-09-935-625-15719
c 28	39	8.4	461	5	US-09-935-625-24218
c 29	39	8.4	461	5	US-09-935-625-27443
c 30	39	8.4	492	5	US-09-620-393B-4294
c 31	38.8	8.3	465	5	US-09-620-393B-9008
c 32	38.8	8.3	1752	5	US-09-886-432-16050
c 33	38.2	8.2	251	5	US-09-975-254-890
c 34	37.8	8.1	479	5	US-09-620-393B-8483
c 35	37.8	8.1	485	5	US-09-620-393B-7814
c 36	37.4	8.0	256	5	US-09-975-254-17947
c 37	37.2	8.0	221	5	US-09-919-002-652
c 38	37	7.9	3083	7	US-10-137-036-1
c 39	37	7.9	3084	7	US-10-137-036-143
c 40	36.8	7.9	399	5	US-09-918-995-7727
c 41	36.6	7.9	239	5	US-09-975-254-1973
c 42	36.6	7.9	265	5	US-09-975-254-13012
c 43	35.8	7.7	234	5	US-09-975-254-19108
c 44	35.8	7.7	261	5	US-09-975-254-18104
c 45	35.6	7.6	552	1	PCT-US02-19457-108

RESULT 1

PCT-US02-18947-1279/c

: Sequence 1279, Application PC/TUS0218947

: GENERAL INFORMATION:

: APPLICANT: Rosetta Inpharmatics

: TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients

: FILE REFERENCE: 9301-175-228

: CURRENT APPLICATION NUMBER: PCT/US02/18947

: CURRENT FILING DATE: 2002-06-14

: PRIOR APPLICATION NUMBER: 60/380,770

: PRIOR FILING DATE: 2002-05-14

: NUMBER OF SEQ ID NOS: 2699

: SEQ ID NO 1279

: LENGTH: 777

: TYPE: DNA

: ORGANISM: Homo sapiens

: PUBLICATION INFORMATION:

: DATABASE ACCESSION NUMBER: NM\_006398

: DATABASE ENTRY DATE: 2001-06-18

PCT-US02-18947-1279

Query Match 85.2%; Score 397; DB 1; Length 777;

Best Local Similarity 98.0%; Pred. No. 4.3e-111;

Matches 444; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

Oy	1	atcaagaacacatagatttcgggcaatactctcatctccatcccccaccccaaatetta	60
Db	631	ATCAAGAACAATAGAGTT-GGGCAATATACTCTACCCATCCACCCAAATCTTA	633
Oy	61	ctctactcatctctcatcttaattttgggaaatcatcagaagatgttcttgagta	120
Db	632	CTCTACTCATCTCTCTCATTAATTTTGGAAATCATCAGAAGATGTTTCGTTGAGTA	573
Oy	121	agaagattaaagaataagctttttgaccccttgcacacaccccccaggggtgttcac	180
Db	572	AGAGATTAAAGAATAATANGCTTTTIGACCCCTGCCAACACCCATCCCGGGTGTAC	513
Oy	161	cttccatacaatadacatgccaggagagtaagtgtccctttctgagtcgaataatcgc	240
Db	512	CTCCCAATACAATAGATGCCAGGAGAGTAAGTTGGCCCTTCTGTATGCCGTATCTGCC	453
Oy	241	atcatcttcccattctccagtcctcttccattgcaagtcacaatcttggttcaggat	300
Db	452	ATCATCTTCCCAATCTCCAGTCT-CTTCCATTGCAAGTCACAAATCTGGGTCTCAGGAT	394
Oy	301	tatacccgcttagtctcgatcattgtttcactgtgccactgagtcgagaccttcgac	360



RESULT . 6  
US-10-040-862-646/c  
: Sequence 646; Application US/10040862  
: GENERAL INFORMATION:  
: APPLICANT: Gaiger, Alexander  
: APPLICANT: Algate, Paul A.  
: APPLICANT: Mannion, Jane

```

RESULT      5
US-10-040-862-205/c
: Sequence 205, Application US/10040862
: GENERAL INFORMATION:
: APPLICANT: Gaiger, Alexander
: APPLICANT: Algate, Paul A.
: APPLICANT: Mannion, Jane
: APPLICANT: Better, Marc
: APPLICANT: Corixa Corporation
: TITLE OF INVENTION: Compositions and
: TITLE OF INVENTION: Hematological M
: FILE REFERENCE: 014058-013520US
: CURRENT APPLICATION NUMBER: US/10/04
: CURRENT FILING DATE: 2001-11-06
: PRIOR APPLICATION NUMBER: US 60/186,
: PRIOR FILING DATE: 2000-03-01

```

```

APPLICANT: Retter, Marc
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
of Hematological Malignancies
FILE REFERENCE: 014058-013520US
CURRENT APPLICATION NUMBER: US/10/040,862
CURRENT FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: US 60/223,378
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: US 09/796,692
PRIOR FILING DATE: 2001-03-01
NUMBER OF SEQ ID NOS: 10467
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 646
LENGTH: 474
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)..(474)
OTHER INFORMATION: n = A,T,C or G
US-10-040-862-646

Query Match 71.8%; Score 334.4; DB 7; Length 474;
Best Local Similarity 99.2%; Pred. No. 4.9e-92;
Matches 357; Conservative 1; Mismatches 1; Indels 2; Gaps 2:

Qy 1 atcaagaacatagagtcgagctaacctatctctaccctaccctcccccacccaattcta 60
Db 358 ATCAAGAACAATAGAGTT-GGGCAATATCTTCATCTACCCATCCCAACCAATCTTA 300
Qy 61 ctctactcatctcattctcatttaattttgggaaatcacaagaatgtgttcgttgagta 120
Db 299 CTCTACTCATCTCATTCATTAATTTGGGAAATCATCAGAAGATGTGTCGTTGAGTA 240
Qy 121 agagattaaaagaataaagctttttgacccctgcacaacccccatgccagggtggtcac 180
Db 239 AGAGATTAAAGAATAAAGCTTTTGATCCCTGCCAACACCCCATGCCAGGGTGGTCAC 180
Qy 181 cctccaatacaataacatgcccaggaagatgaagtgcctcttctgatccgtaactctacc 240
Db 179 CCTCCAATACAATAACATGCAGGAAGATGAAGTTGCCCTTTCTGAITGGCGTAATCTGCC 120
Qy 241 atcattctcccatctccagctctcctttccattgcgaagtcaacaattcggtctcagggat 300
Db 119 ATCATCTTCCCATCTTCCAGTCT-CITTCATTGCAAGTCACAATCTGGGTCTCAGGAT 61
Qy 301 tatacccgcttagctcgtatcattgtctttcacttgtgcacttgacgtgaccttcgcac 360
Db 60 TATACCCGCTTAGCTCGATCATGTCTTCACTGTGGCACTGAGCTGGACCTTCGCAC 1

```

Db 239 AGAGATTAAAGAAATAAGCTTTTGGATCCCTGCGCAACACCCCATGCCAGGGTGGTCAC 180  
Qy 181 cctcccaataacaatccaggaagagtgagttggcccttcttgatgccgtaatactgccc 240  
Db 179 CTCCCAATACATAACATCCAGGAAGAGTAAAGTTGCCCTTTCTGATGCCGTAATCTGCC 120  
Qy 241 atcatcttcccatcttcagctctcttccatgcagatcacaactcgggtctcaaggat 300  
Db 119 ATATCTCTCCCATCTTCCAGTCT-CTTCCATGCGAAGTCAACAATCTGGGTCACAGGAT 61  
Qy 301 tataccgctttagctcgatcattcttccatcttgcacttgccactgagctggaccttcgac 360  
Db 60 TATACCGCTTGTAGTCGATCATCTTTCACCTTGTGCGCATGAGTGGACCTTCGCAC 1

RESULT 8  
US-10-040-862-248/c  
; Sequence 248, Application US/10040862  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013520US  
; CURRENT APPLICATION NUMBER: US/10/040,862  
; CURRENT FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: US 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; PRIOR APPLICATION NUMBER: US 09/796,692  
; PRIOR FILING DATE: 2001-03-01  
; NUMBER OF SEQ ID NOS: 10467  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 248  
; LENGTH: 453  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(453)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-040-862-248

Query Match 69.4%; Score 323.4; DB 7; Length 453;  
Best Local Similarity 99.1%; Pred. No. 1.1e-88;  
Matches 346; Conservative, 0; Mismatches 1; Indels 2; Gaps 2;  
Qy 1 atcaagaatacagatagttcgggagataatacttctcaccatcccccatacttta 60

Db 347 ATCAAGAAACATAGAGTT-GGCAATATACCTTCACTCCCTACCCATCCCAAAATCTTA 289  
Qy 61 cttactcatctctcattcatttgggaaatcatcagaagatgtgtcgttgagta 120  
Db 288 CTCTACTCACTCTCACTCTCACTTAAATTTGGGAATCATCAGAAGATGTGTTCTGTTGAGTA 229  
Qy 121 agattataaagaataagctttttgacccttgcacacaccccatgcccagggatgacac 180  
Db 228 AGAGATTAAAGAAATAAGCTTTTGGACCTCTGCCAACACCCCATGCCAGGGTGGTCAC 169  
Qy 181 cctcccaataacaatccaggaagagtgagttggcccttcttgatgccgtaatactgccc 240  
Db 168 CTTCATATACATTAAGATGCGAGGAAGTAAAGTTGCCCTTTCTGATGCCGTAATCTGCC 109  
Qy 241 atcatcttcccatcttcagctctcttccatcttgcacttgccactgagctggatcaggat 300  
Db 108 ATCATCTTCCCATCTTCCAGTCT-CTTTCATTTCAAGTCACAATCTGGGTCACAGGAT 50  
Qy 301 tataccgctttagctcgatcattcttccatcttgcacttgccactgagctg 349  
Db 49 TATACCGCTTGTAGTCGATCATCTTTCACCTTGTGCGCATGAGCTG 1

RESULT 9  
US-10-040-862-5273/c  
; Sequence 5273, Application US/10040862  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther  
; FILE REFERENCE: 014058-013520US  
; CURRENT APPLICATION NUMBER: US/10/040,862  
; CURRENT FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: US 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; PRIOR APPLICATION NUMBER: US 09/796,692  
; PRIOR FILING DATE: 2001-03-01  
; NUMBER OF SEQ ID NOS: 10467  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5273  
; LENGTH: 453  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (420)

; OTHER INFORMATION: n=A,T,C or G  
 US-10-040-862-5273

Query Match 69.4%; Score 323.4; DB 7; Length 453;  
Best Local Similarity 99.1%; Pred. NO. 1.1e-88;  
Matches 346. Conservative 0; Mismatches 1; Indels 2

Qy	1	atcaagaacatagagttcgggcaatatctctcattccatcccatccaccaccaactctta	60
Db	347	ATCAAGAACAATAGAGT-TGGGCAATATACITTCATCTACCATCCACCACCAATCTTA	289
Qy	61	ctctactcatctatctcatatttttgggaaatcatcagaagatgtgttcgttgagta	120
Db	288	CTCTACTCATCTCATTTCTCATTAATTTTGGGAATCATCAGAAGATGTGTTTCGTTGAGTA	229
Qy	121	agagattaaagaataaagctttttgacccttgccaaacacccatgccagggtgggtcac	180
Db	228	AGAGATTAAAAGAAATTAAGCTTTTTCACCCCTTGCCNACACCCCATGCCAGGGTGGTTCAC	169
Qy	181	ctccaataacaataacatgccaggaagagtaagttgcccctttctgatgcgtaatatgcc	240
Db	168	CCTCCAATACAAATAAGATGCCAGGAAGAGTAAGTTGCCCTTTCTGATGCCGTAATCTGCC	109
Qy	241	atcatcttcccatcttccagttcccttccattcgaagtcaaatctgggtctcaggagat	300
Db	108	ATCATCTTCCCACTCTTCAGTCT-CTTTCATTCGAAGTCAACATCTGGGTCTCAGGGAT	50
Qy	301	tataccgctttagctcgcatactgttcttcactgtgccaactgagctg	349
Db	49	TATACCGGTCITAGTCTCGATCATTTGCTTTCATCTTGCCACTGAGCTG	1

RESULT 10  
US-10-137-036-34/c  
: Sequence 34, Application US/10137036

```

: GENERAL INFORMATION:
: APPLICANT: Perera, Ranjan
: APPLICANT: Rice, Stephen
: APPLICANT: Eagleton, Clare
: APPLICANT: Lasham, Annette
: APPLICANT: Wood, Marion
: APPLICANT: Visser, Elizabeth
: TITLE OF INVENTION: Compositions and Methods for the
: TITLE OF INVENTION: Modification of Gene Expression
: FILE REFERENCE: 11000-1036c4
: CURRENT APPLICATION NUMBER: US/10/137,036
: CURRENT FILING DATE: 2002-04-30
: PRIORITY APPLICATION NUMBER: PCT/NZ 01/00115
: PRIORITY FILING DATE: 2001-06-20
: PRIORITY APPLICATION NUMBER: U.S. No. 09/724,624
: PRIORITY FILING DATE: 2000-11-28
: PRIORITY APPLICATION NUMBER: U.S. No. 09/598,401
: PRIORITY FILING DATE: 2000-06-20
: PRIORITY APPLICATION NUMBER: PCT/NZ00/00018
: PRIORITY FILING DATE: 2000-02-24
: PRIORITY APPLICATION NUMBER: U.S. No. 60/146,591
: PRIORITY FILING DATE: 1999-07-30
: PRIORITY APPLICATION NUMBER: U.S. No. 09/276,599
: PRIORITY FILING DATE: 1999-03-25
: NUMBER OF SEQ ID NOS: 143
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 34
: LENGTH: 1274
: TYPE: DNA
: ORGANISM: Eucalyptus grandis
: US-10-137-036-34

```

Query Match 9.8%; Score 45.8; DB 7; Length 1274;  
Best Local Similarity 57.8%; Pred. No. 0.0011;  
Matches 100: Conservative 0; Mismatches 72; Indels 1; Gaps 1;

```

Qy 290 ccagggaagtaagtgcctttctgatgcgcgtaataatcgccatcatcttcccatctcca 259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1241 CCAAGTGGAGGGTGGACTCCTTCITGAAIGTTATTAATCGGCCAAGGTTGGCCCACTTCCA 1182
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 260 gtctcctttccattgcaagtcacaatcgggtctcagggtattataccgcgtcttagtctcg 319
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1181 G-CTGCTTGGCAGCAAGATAAGCCTCTGCTGGTCCGAGGGATCCCTTCTCTGTCTCTGG 1123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 320 atcatcttttcattctgtgcactgagctgagcacttcgcacacctgggaggagggt 372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1122 ATTTTGTCTTCACATTATCAACTGTGTCGAGGCTTCCACCTCGAGGGTGAT 1070

RESULT 11
US-09-620-393B-2769/c
: Sequence 2769, Application US/09620393B
: GENERAL INFORMATION:
: APPLICANT: ALEXANDROV, Nikolai et al.
: TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
: TITLE OF INVENTION: THEREBY
: FILE REFERENCE: 2750-1068P
: CURRENT APPLICATION NUMBER: US/09/620,393B
: CURRENT FILING DATE: 2000-07-21
: NUMBER OF SEQ ID NOS: 9948
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 2769
: LENGTH: 442
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
: FEATURES:
: NAME/KEY: misc_feature
: LOCATION: 1..442
: OTHER INFORMATION: any n = a, g, c, t, unknown, or other
: FEATURES:
: NAME/KEY: misc_feature
: LOCATION: 1..442
: OTHER INFORMATION: Ceres Seq. ID 1384117
US-09-620-393B-2769

```

[illegible]

```

RESULT 12
US-09-620-393B-4443/c
: Sequence 4443, Application US/09620393B
: GENERAL INFORMATION:
: APPLICANT: ALEXANDROV, Nikolai et al.
: TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
: TITLE OF INVENTION: THEREBY
: FILE REFERENCE: 2750-1068P
: CURRENT APPLICATION NUMBER: US/09/620,393B
: CURRENT FILING DATE: 2000-07-21
: NUMBER OF SEQ ID NOS: 9948
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 4443
: LENGTH: 501
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana

```

```

: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1..501
: OTHER INFORMATION: any n = a, g, c, t, unknown, or other
:
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1..501
: OTHER INFORMATION: Ceres Seq. ID 1388198
:
: US-09-620-393B-4443

Query Match          9.5%; Score 44.2; DB 5; Length 501;
Best Local Similarity 57.2%; Pred. NO. 0.0024;
Matches 99; Conservative .0; Mismatches 73; Indels 1; Gaps 1;

QY 200 ccaggagagatgaattgcccctttctgatgccgtaattgccatcatcttccccattccca 259
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 274 CCAGTGGAGGGTGGATTCCTTCTGGATATGTGAATCAGCCACAGGTACGCCCACTCTCTA 215
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    250 cctctctcttccctgcagctccagctatccagctcagagattatccattctattctg 119

```

```
Db      QY      Db      214 G-CTCCTTCGCGGGAAAATAAGCCCTTGCTGTATCCGAGGAATGCCCTTCTTAACCTGG   156
                320 atcattgttttcaccttgctgcacctaagcttgcaccctcggaaggagt    372
                    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      155 ATCTTGGGCCTTAACGTTGTGATGTTGTCGGAGCTTTCACCCTCGAAGGTGAT    103

RESULT          13
US-09-975-254-22419/c
; Sequence 22419, Application US/09975254
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
```

```

: CURRENT APPLICATION NUMBER: US/09/975,254
: CURRENT FILING DATE: 2001-10-12
: PRIOR APPLICATION NUMBER: US/09/263,191
: PRIOR FILING DATE: 1999-03-05
: NUMBER OF SEQ ID NOS: 31255
: SEQ ID NO 22419
: LENGTH: 243
: TYPE: DNA
: ORGANISM: Glycine max
: OTHER INFORMATION: Clone ID: 700960992H1
US-09-975-254-22419

Query Match          9.3%; Score 43.4; DB 5; Length 243:
Best Local Similarity 57.4%; Pred. No. 0.0032;
Matches 97; Conservative 0; Mismatches 71; Indels 1; Gaps 1:

Qy 218 cttctctgatcgcgtaattgccatcatcttcccacattctccagtcctcttccattgcaa 277
      |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 213 CCTTCTGGAATGTTAATACAGCAGGGTCTTCTCAATCTCAAG-CTGCTTCGGGCAAG 155
      |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Qy 278 gtcaaatctgggtctcagggaattataccgctcttagctctgacatctgctttccactgt 337
      |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 154 ATAAGCCTCTGCTGGTCTGGGGGAATACCCCTCCCTTATCCTGAACTCTGGCTTTTCACATTA 95
      |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

```

Db 94 TCAATTCATCAGAACTTTCCACCTCCAAGGTGATGGTCTTTTCCGTGCA 46

RESULT 14

US-09-620-393B-1362/c

: Sequence 1362, Application US/09620393B

: GENERAL INFORMATION:



Qy 278 gtcacaatctgggtctcagggtattacccgtctttagtctcgcgatcatgctttcacttgt 337  
| | |||| | | |||| | | |||| | | |||| | | ||||  
Db 197 ATAAGCCTCTGCTGATCCGGAGGAATACCTTCCTTATCCTGGATCTTGGCCTTAACGTTG 138  
Qy 338 gccactgagctgaccttcgcacccctgggagaggt 372  
| | | | | | |||| | | ||||  
Db 137 TCGATGGTGCAGAGCTTCCACCTCGAGGGTGAT 103

Search completed: July 29, 2002, 19:27:19  
Job time: 7610 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 16:37:39 ; Search time 4748.22 Seconds  
(without alignments)  
2123.293 Million cell updates/sec

Title: US-09-880-107-2492  
Perfect score: 466  
Sequence: 1 atcaagaacatagattc.....ggtaagggtggatggtctt 466

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_NA\_Main.\*  
1: /cgn2\_6/ptodata/2/pna/PCTUS\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/pna/US06\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/pna/US07\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/pna/US0801\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/pna/US082\_COMB.seq.\*  
7: /cgn2\_6/ptodata/2/pna/US083\_COMB.seq.\*  
8: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq.\*  
9: /cgn2\_6/ptodata/2/pna/US085\_COMB.seq.\*  
10: /cgn2\_6/ptodata/2/pna/US086\_COMB.seq.\*  
11: /cgn2\_6/ptodata/2/pna/US087\_COMB.seq.\*  
12: /cgn2\_6/ptodata/2/pna/US088\_COMB.seq.\*  
13: /cgn2\_6/ptodata/2/pna/US089\_COMB.seq.\*  
14: /cgn2\_6/ptodata/2/pna/US090\_COMB.seq.\*  
15: /cgn2\_6/ptodata/2/pna/US091\_COMB.seq.\*  
16: /cgn2\_6/ptodata/2/pna/US092\_COMB.seq.\*  
17: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq.\*  
18: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq.\*  
19: /cgn2\_6/ptodata/2/pna/US095A\_COMB.seq.\*  
20: /cgn2\_6/ptodata/2/pna/US095B\_COMB.seq.\*  
21: /cgn2\_6/ptodata/2/pna/US095C\_COMB.seq.\*  
22: /cgn2\_6/ptodata/2/pna/US095D\_COMB.seq.\*  
23: /cgn2\_6/ptodata/2/pna/US096A\_COMB.seq.\*  
24: /cgn2\_6/ptodata/2/pna/US096B\_COMB.seq.\*  
25: /cgn2\_6/ptodata/2/pna/US096C\_COMB.seq.\*  
26: /cgn2\_6/ptodata/2/pna/US096D\_COMB.seq.\*  
27: /cgn2\_6/ptodata/2/pna/US096E\_COMB.seq.\*  
28: /cgn2\_6/ptodata/2/pna/US097A\_COMB.seq.\*  
29: /cgn2\_6/ptodata/2/pna/US097B\_COMB.seq.\*  
30: /cgn2\_6/ptodata/2/pna/US097C\_COMB.seq.\*  
31: /cgn2\_6/ptodata/2/pna/US098A\_COMB.seq.\*  
32: /cgn2\_6/ptodata/2/pna/US098B\_COMB.seq.\*  
33: /cgn2\_6/ptodata/2/pna/US098C\_COMB.seq.\*  
34: /cgn2\_6/ptodata/2/pna/US099A\_COMB.seq.\*  
35: /cgn2\_6/ptodata/2/pna/US099B\_COMB.seq.\*  
36: /cgn2\_6/ptodata/2/pna/US099C\_COMB.seq.\*  
37: /cgn2\_6/ptodata/2/pna/US100\_COMB.seq.\*  
38: /cgn2\_6/ptodata/2/pna/US101\_COMB.seq.\*  
39: /cgn2\_6/ptodata/2/pna/US6000\_COMB.seq.\*  
40: /cgn2\_6/ptodata/2/pna/US6001\_COMB.seq.\*  
41: /cgn2\_6/ptodata/2/pna/US6002\_COMB.seq.\*  
42: /cgn2\_6/ptodata/2/pna/US6003\_COMB.seq.\*  
43: /cgn2\_6/ptodata/2/pna/US6004\_COMB.seq.\*

44: /cgn2\_6/ptodata/2/pna/US6005\_COMB.seq.\*  
45: /cgn2\_6/ptodata/2/pna/US6006\_COMB.seq.\*  
46: /cgn2\_6/ptodata/2/pna/US6007\_COMB.seq.\*  
47: /cgn2\_6/ptodata/2/pna/US6008\_COMB.seq.\*  
48: /cgn2\_6/ptodata/2/pna/US6009\_COMB.seq.\*  
49: /cgn2\_6/ptodata/2/pna/US6010\_COMB.seq.\*  
50: /cgn2\_6/ptodata/2/pna/US6011\_COMB.seq.\*  
51: /cgn2\_6/ptodata/2/pna/US6012\_COMB.seq.\*  
52: /cgn2\_6/ptodata/2/pna/US6013\_COMB.seq.\*  
53: /cgn2\_6/ptodata/2/pna/US6014\_COMB.seq.\*  
54: /cgn2\_6/ptodata/2/pna/US6015\_COMB.seq.\*  
55: /cgn2\_6/ptodata/2/pna/US6016\_COMB.seq.\*  
56: /cgn2\_6/ptodata/2/pna/US6017\_COMB.seq.\*  
57: /cgn2\_6/ptodata/2/pna/US6018\_COMB.seq.\*  
58: /cgn2\_6/ptodata/2/pna/US6019\_COMB.seq.\*  
59: /cgn2\_6/ptodata/2/pna/US6020\_COMB.seq.\*  
60: /cgn2\_6/ptodata/2/pna/US6021\_COMB.seq.\*  
61: /cgn2\_6/ptodata/2/pna/US6022\_COMB.seq.\*  
62: /cgn2\_6/ptodata/2/pna/US6023\_COMB.seq.\*  
63: /cgn2\_6/ptodata/2/pna/US6024\_COMB.seq.\*  
64: /cgn2\_6/ptodata/2/pna/US6025\_COMB.seq.\*  
65: /cgn2\_6/ptodata/2/pna/US6026\_COMB.seq.\*  
66: /cgn2\_6/ptodata/2/pna/US6027\_COMB.seq.\*  
67: /cgn2\_6/ptodata/2/pna/US6028\_COMB.seq.\*  
68: /cgn2\_6/ptodata/2/pna/US6029\_COMB.seq.\*  
69: /cgn2\_6/ptodata/2/pna/US6030\_COMB.seq.\*  
70: /cgn2\_6/ptodata/2/pna/US6031\_COMB.seq.\*  
71: /cgn2\_6/ptodata/2/pna/US6032\_COMB.seq.\*  
72: /cgn2\_6/ptodata/2/pna/US6033\_COMB.seq.\*  
73: /cgn2\_6/ptodata/2/pna/US6034\_COMB.seq.\*  
74: /cgn2\_6/ptodata/2/pna/US6035\_COMB.seq.\*  
75: /cgn2\_6/ptodata/2/pna/US6036\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	IDB	ID	Description
1	465	100.0	466	33	US-09-880-107-2492	Sequence 2492, App
2	430.2	85.9	7315	63	US-60-248-505-445	Sequence 445, App
3	430.2	85.9	9250	63	US-60-248-505-278	Sequence 278, App
4	430.2	85.9	125235	63	US-60-245-225-89	Sequence 89, Appl
5	399.8	85.8	833	1	PCT-US00-05882-475	Sequence 475, App
6	399.8	85.8	833	34	US-09-925-301-475	Sequence 475, App
7	398.6	85.5	798	37	US-10-044-090-482	Sequence 482, App
8	398.6	85.5	798	61	US-60-222-113-231	Sequence 231, App
9	398.6	85.5	798	65	US-60-260-483-482	Sequence 482, App
10	398.6	85.5	951	56	US-60-172-373-12767	Sequence 12767, A
11	398.6	85.5	2697	59	US-60-209-009-485	Sequence 485, App
12	398.6	85.5	2697	59	US-60-209-106-423	Sequence 423, App
13	398.6	85.5	2697	60	US-60-213-360-7551	Sequence 7551, Ap
14	398.6	85.5	2697	66	US-60-278-258-16025	Sequence 16025, A
15	397	85.2	777	13	US-08-978-289-1	Sequence 1, Appl1
16	397	85.2	777	36	US-09-994-444-1	Sequence 4218, Ap
17	397	85.2	806	23	US-09-606-680-4218	Sequence 98, Appl
18	397	85.2	32659	59	US-60-206-020-98	Sequence 1379, Ap
19	350.2	83.7	797	18	US-09-471-275-1379	Sequence 10199, A
20	379	81.3	588	37	US-10-029-386-10199	Sequence 12723, A
21	379	81.3	588	37	US-10-029-386-12723	Sequence 205, App
22	334.4	71.8	474	30	US-09-796-692-205	Sequence 646, App
23	334.4	71.8	474	30	US-09-796-692-646	Sequence 4647, Ap
24	334.4	71.8	474	30	US-09-796-692-4647	Sequence 6673, Ap
25	339.4	70.7	568	25	US-09-644-870-6673	Sequence 10267, A
26	325	69.7	384	18	US-09-431-517-10267	Sequence 248, App
27	323.4	69.4	453	30	US-09-796-692-248	Sequence 5273, Ap
28	323.4	69.4	453	30	US-09-796-692-5273	Sequence 219, App
29	316.2	67.9	583	55	US-60-164-378-219	Sequence 3905, Ap
30	316.2	67.9	371	15	US-09-179-473-3905	Sequence 3905, Ap
31	316.2	67.9	371	17	US-09-328-351-3905	Sequence 3905, Ap

RESULT 3  
US-60-248-505-278  
: Sequence 278, Application US/60248505  
: GENERAL INFORMATION:  
: APPLICANT: Beasley, Ellen  
: TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED

; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
; FILE REFERENCE: C1000918  
; CURRENT APPLICATION NUMBER: US/60/248,505  
; CURRENT FILING DATE: 2000-11-15  
; NUMBER OF SEQ ID NOS: 1998  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 278  
; LENGTH: 9250  
; TYPE: DNA  
; ORGANISM: human  
US-60-248-505-278

Query Match 85.9%; Score 400.2; DB 63; Length 9250;  
Best Local Similarity 98.5%; Pred. No. 2.9e-110;  
Matches 446; Conservative 0; Mismatches 3; Indels 4; Gaps 4;  
QY 1 atcaagaacatagagtcgggcaataatacttcatctaccatccaccacccaatctta 60  
DB 8790 atcaagaacatagagtt-gggcaataatacttcatctaccatccaccacccaatctta 8848  
QY 61 ctctactcatctctcatttaatttttgggaataatcatcagaagaatgtgttcgttgagta 120  
DB 8849 ctctactcatctctcatttaatttttgggaataatcatcagaagaatgtgttcgttgagta 8908  
QY 121 agagattaaaagaataagctttttgacccctgccacaccccatgccagggtggtcac 180  
DB 8909 agagattaaaagaataagctttttgacccctgccacaccccatgccagggtggtcac 8968  
QY 181 ctcccaatacaataatgcttttgcctcctccttccattgcaagtcacaatctggttcaggat 240  
DB 8969 ctcccaatacaataatgcttttgcctcctccttccattgcaagtcacaatctggttcaggat 9028  
QY 241 atcatcttcccatcttcagctctcttccattgcaagtcacaatctggttcaggat 300  
DB 9029 atcatcttcccatcttcagctctcttccattgcaagtcacaatctggttcaggat 9087  
QY 301 tatacccgcttagtctcgatcattgtttcacttgccactgagctggacccttcgcac 360  
DB 9088 tatacccgcttagtctcgatcattgtttcacttgccactgagctggacccttcgcac 9147  
QY 361 ctggaggagagtgcttttgcctcctcctcctcctcctcctcctcctcctcctcct 420  
DB 9148 ct-ggaggagggtgcttttgcctcctcctcctcctcctcctcctcctcctcctcct 9205  
QY 421 catcactggggttcaccactttcagggttaag 453  
DB 9207 catcact-ggggttcaccactttcagggttaag 9238

RESULT 4  
US-60-245-225-89  
; Sequence 89, Application US/60245225  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
; RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
; FILE REFERENCE: C1000885  
; CURRENT APPLICATION NUMBER: US/60/245,225  
; CURRENT FILING DATE: 2000-11-03  
; NUMBER OF SEQ ID NOS: 705  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 89  
; LENGTH: 125235  
; TYPE: DNA  
; ORGANISM: Human  
US-60-245-225-89

Query Match 85.9%; Score 400.2; DB 63; Length 125235;  
Best Local Similarity 98.5%; Pred. No. 8.9e-110;

Matches 446; Conservative 0; Mismatches 3; Indels 4; Gaps 4;  
QY 1 atcaagaacatagagtcgggcaataatacttcatctaccatccaccacccaatctta 60  
DB 95978 atcaagaacatagagtt-gggcaataatacttcatctaccatccaccacccaatctta 96036  
QY 61 ctctactcatctcattcatttaatttttgggaataatcatcagaagaatgtgttcgttgagta 120  
DB 96037 ctctactcatctcattcatttaatttttgggaataatcatcagaagaatgtgttcgttgagta 96096  
QY 121 agagattaaaagaataagctttttgacccctgccacaccccatgccagggtggtcac 180  
DB 96097 agagattaaaagaataagctttttgacccctgccacaccccatgccagggtggtcac 96156  
QY 181 ctcccaatacaataatgcttttgcctcctccttccattgcaagtcacaatctggttcaggat 240  
DB 96157 ctcccaatacaataatgcttttgcctcctccttccattgcaagtcacaatctggttcaggat 96216  
QY 241 atcatcttcccatcttcagctctcttccattgcaagtcacaatctggttcaggat 300  
DB 96217 atcatcttcccatcttcagctctcttccattgcaagtcacaatctggttcaggat 96275  
QY 301 tatacccgcttagtctcgatcattgtttcacttgccactgagctggacccttcgcac 360  
DB 96276 tatacccgcttagtctcgatcattgtttcacttgccactgagctggacccttcgcac 96335  
QY 361 ctggaggagagtgcttttgcctcctcctcctcctcctcctcctcctcctcctcct 420  
DB 96336 ct-ggaggagggtgcttttgcctcctcctcctcctcctcctcctcctcctcctcct 96394  
QY 421 catcactggggttcaccactttcagggttaag 453  
DB 96395 catcact-ggggttcaccactttcagggttaag 96426

RESULT 5  
PCT-US00-05882-475/c  
; Sequence 475, Application PC/TUS0005882  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; APPLICANT: Steve Ruben  
; TITLE OF INVENTION: Human Cancer Associated Gene Sequences and Polypeptides  
; FILE REFERENCE: P4106PCT  
; CURRENT APPLICATION NUMBER: PCT/US00/05882  
; CURRENT FILING DATE: 2000-03-08  
; EARLIER FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1694  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 475  
; LENGTH: 833  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (9)  
; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (15)  
; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (29)  
; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (58)  
; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (73)  
; OTHER INFORMATION: n equals a,t,g, or c

PCT-US00-05882-475

Query Match 85.8%; Score 399.8; DB 1; Length 833;  
Best Local Similarity 98.2%; Pred. No. 1.4e-110;  
Matches 445; Conservative 1; Mismatches 3; Indels 4; Gaps 4;  
Qy 1 atcaagaacacatagagtcgggcaataatacttcatctcaccatccaccacaaatctta 60  
Db 751 ATCAAGAACAATAGAGTT-GGGCAATATACTTCTATCTACCCATCCCAACAAATCTTA 693  
Qy 61 ctctactcatctctcatcttaattttgggaaatcatcagaagatgtgtcgttgagta 120  
Db 692 CTCTACTCATCTCTCATTAATTTTGGGAAATCATCAGAAGATGTGTTCGTGAGTA 633  
Qy 121 agagattaaagaataagctttttgacctgccaacaccccatgccagggtgggtcac 180  
Db 632 AGAGATTAAAGAAATAAGCTTTTGACCCCTGCCAACCCCAATGCCAGGGTGGTCAC 573  
Qy 181 ctccaatacaataacatgccagggaagtaagttggcccttctctgatgccgtaactcgtcc 240  
Db 572 CTCCAATAACAATAASATGCCAGGAAGTAAGTTGCCCTTCTGTATGCCGTAATCTGCC 513  
Qy 241 atcatcttccatctccagtcctcttccatcttccatcttccatcttccatcttccatct 300  
Db 512 ATCATCTTCCCATCTCCAGTCT-CTTTCCATTGCAAGTCACAATCTGGGCTCAGGGAT 454  
Qy 301 tataccgctcttagctcgatcatcttctcacttgccactgagctggaccttcgcac 360  
Db 453 TATACCCGCTTAGTCTCGATCATCTTCTCACTTGTGCCACTGAGCTGGACCTTCGCAC 394  
Qy 361 ctggggagggtgctcttcttgcctcatcactgactccacaagaagaaggagcactcct 420  
Db 393 CT-GGAGGAGGTGCTCTTTTGCTCATCACTGACTTCCACAAAGAAACAAGGCGACTCT 335  
Qy 421 catcactggggttccaccatttcagggtgaag 453  
Db 334 CATCACT-GGGCTTCACCACTTTCAGGGTAAGG 303

RESULT 6

US-09-925-301-475/c  
; Sequence 475; Application US/09925301  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA106  
; CURRENT APPLICATION NUMBER: US/09/925,301  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05882  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1694  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 475  
; LENGTH: 833  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (9)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (15)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (29)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (58)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature

LOCATION: (73)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-301-475

Query Match 85.8%; Score 399.8; DB 34; Length 833;  
Best Local Similarity 98.2%; Pred. No. 1.4e-110;  
Matches 445; Conservative 1; Mismatches 3; Indels 4; Gaps 4;  
Qy 1 atcaagaacacatagagtcgggcaataatacttcatctcaccatccaccacaaatctta 60  
Db 751 ATCAAGAACAATAGAGTT-GGGCAATATACTTCTATCTACCCATCCCAACAAATCTTA 693  
Qy 61 ctctactcatctctcatcttaattttgggaaatcatcagaagatgtgtcgttgagta 120  
Db 692 CTCTACTCATCTCTCATTAATTTTGGGAAATCATCAGAAGATGTGTTCGTGAGTA 633  
Qy 121 agagattaaagaataagctttttgacctgccaacaccccatgccagggtgggtcac 180  
Db 632 AGAGATTAAAGAAATAAGCTTTTGACCCCTGCCAACCCCAATGCCAGGGTGGTCAC 573  
Qy 181 ctccaatacaataacatgccagggaagtaagttggcccttctctgatgccgtaactcgtcc 240  
Db 572 CTCCAATAACAATAASATGCCAGGAAGTAAGTTGCCCTTCTGTATGCCGTAATCTGCC 513  
Qy 241 atcatcttccatctccagtcctcttccatcttccatcttccatcttccatcttccatct 300  
Db 512 ATCATCTTCCCATCTCCAGTCT-CTTTCCATTGCAAGTCACAATCTGGGCTCAGGGAT 454  
Qy 301 tataccgctttagctcgatcatcttctcacttgccactgagctggaccttcgcac 360  
Db 453 TATACCCGCTTAGTCTCGATCATCTTCTCACTTGTGCCACTGAGCTGGACCTTCGCAC 394  
Qy 361 ctggggagggtgctcttcttgcctcatcactgactccacaagaagaaggagcactcct 420  
Db 393 CT-GGAGGAGGTGCTCTTTTGCTCATCACTGACTTCCACAAAGAAACAAGGCGACTCT 335  
Qy 421 catcactggggttccaccatttcagggtgaag 453  
Db 334 CATCACT-GGGCTTCACCACTTTCAGGGTAAGG 303

RESULT 7

US-10-044-090-482/c  
; Sequence 482; Application US/10044090  
; GENERAL INFORMATION:  
; APPLICANT: Olga Bandman  
; TITLE OF INVENTION: GENES DIFFERENTIALY EXPRESSED IN VASCULAR TISSUE ACTIVATION  
; FILE REFERENCE: PA-0028 US  
; CURRENT APPLICATION NUMBER: US/10/044,090  
; CURRENT FILING DATE: 2002-01-09  
; NUMBER OF SEQ ID NOS: 850  
; SOFTWARE: PERL Program  
; SEQ ID NO 482  
; LENGTH: 798  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No: 575983CB1  
US-10-044-090-482

Query Match 85.5%; Score 398.6; DB 37; Length 798;  
Best Local Similarity 98.2%; Pred. No. 3.1e-110;  
Matches 445; Conservative 0; Mismatches 4; Indels 4; Gaps 4;  
Qy 1 atcaagaacacatagagtcgggcaataatacttcatctcaccatccaccacaaatctta 60  
Db 709 ATCAAGAACAATAGAGTT-GGGCAATATACTTCTATCTACCCATCCCAACAAATCTTA 651  
Qy 61 ctctactcatctctcatcttaattttgggaaatcatcagaagatgtgtcgttgagta 120

```
Db 650 CTCTACTCATCTCATCTCAATTAATTTGGGAATCATCAGAAATGTTGCTTGAGTA 591
Qy 121 agagattaaagaataagctttttgacccctgccaacacccatgccagggtggtcac 180
Db 590 AGAGATTAAGAATAAGCTTTTGAACCCCTGCCAACACCCATGCCCAGGGTGTAC 531
Qy 181 cttcaatacaataacatccaggaagatgaattggcctttctgtatccgtaactgcc 240
Db 530 CTCCAATAACAATAGATGCCAGGAAGAGTAAGTTGCCCTTTCTGATGCCGTAATCTGCC 471
Qy 241 atcatcttcccatcttccagtcctcttccattgcaagtccaactctgggtctcaggat 300
Db 470 ATCATCTTCCCATCTTCCAGTCT-CTTTCCATTGCAAGTCAACAATCTGGGTCACGGAT 412
Qy 301 tataccgcttagtcgatcattgtcttcaactgtgccaactgccaggtggtcac 360
Db 411 TATACCCGCTTAGTCTCGATCATCTGCTTTCACCTTGTGCCACTGAGCTGGACCTTCGCAC 352
Qy 361 ctggaggaggtgctcttctgctcatcactgactcaccagaaacaaaggagcgtcct 420
Db 351 CT-GGAGGAGGTGCTCTTTGCTCTCATCAGTACTCCACAAGAAACAAAGGGCAGCTCCT 293
Qy 421 catcactggggtttcaccactttcagggttaag 453
Db 292 CATCACT-GGGCTTACCACCTTTTCAGGGTAAGG 261

RESULT 8
US-60-222-113-231/c
; Sequence 231, Application US/60222113
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 P
; CURRENT APPLICATION NUMBER: US/60/222,113
; CURRENT FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 231
; LENGTH: 798
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 575983CB1
US-60-222-113-231

Query Match 85.5%; Score 398.6; DB 61; Length 798;
Best Local Similarity 98.2%; Pred. No. 3.1e-110;
Matches 445; Conservative 0; Mismatches 4; Indels 4; Gaps 4;

Qy 1 atcaagaacatagagttcgggcaatacttcatctcactccatccaccaccccaaatctta 60
Db 709 ATCAAGAAACATAGAGTT-GGGCAATATATCTTCCATCTACCCATCCCAAAATCTTA 651
Qy 61 ctctactcattctcatttaatttggaaatcatcagaagatgtgtcttgagta 120
Db 650 CTCTACTCATCTCATCTCAATTAATTTGGGAATCATCAGAAATGTTGCTTGAGTA 591
Qy 121 agagattaaagaataagctttttgacccctgccaacacccatgccagggtggtcac 180
Db 590 AGAGATTAAGAATAAGCTTTTGAACCCCTGCCAACACCCATGCCCAGGGTGTAC 531
Qy 181 cttcaatacaataacatccaggaagatgaattggcctttctgtatccgtaactgcc 240
Db 530 CTCCAATAACAATAGATGCCAGGAAGAGTAAGTTGCCCTTTCTGATGCCGTAATCTGCC 471
Qy 241 atcatcttcccatcttccagtcctcttccattgcaagtccaactctgggtctcaggat 300
Db 470 ATCATCTTCCCATCTTCCAGTCT-CTTTCCATTGCAAGTCAACAATCTGGGTCACGGAT 412
Qy 301 tataccgcttagtcgatcattgtcttcaactgtgccaactgccaggtggtcac 360
Db 411 TATACCCGCTTAGTCTCGATCATCTGCTTTCACCTTGTGCCACTGAGCTGGACCTTCGCAC 352
Qy 361 ctggaggaggtgctcttctgctcatcactgactcaccagaaacaaaggagcgtcct 420
Db 351 CT-GGAGGAGGTGCTCTTTGCTCTCATCAGTACTCCACAAGAAACAAAGGGCAGCTCCT 293
Qy 421 catcactggggtttcaccactttcagggttaag 453
Db 292 CATCACT-GGGCTTACCACCTTTTCAGGGTAAGG 261
```

```
Db 411 TATACCCGCTTAGTCTCGATCATCTTTCATCTTGGCCACTGAGCTGGACCTTCGCAC 352
Qy 361 ctggaggaggtgctcttctgctcatcactgactcaccagaaacaaaggagcgtcct 420
Db 351 CT-GGAGGAGGTGCTCTTTGCTCTCATCAGTACTCCACAAGAAACAAAGGGCAGCTCCT 293
Qy 421 catcactggggtttcaccactttcagggttaag 453
Db 292 CATCACT-GGGCTTACCACCTTTTCAGGGTAAGG 261

RESULT 9
US-60-260-483-482/c
; Sequence 482, Application US/60260483
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 P
; CURRENT APPLICATION NUMBER: US/60/260,483
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 482
; LENGTH: 798
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 575983CB1
US-60-260-483-482

Query Match 85.5%; Score 398.6; DB 65; Length 798;
Best Local Similarity 98.2%; Pred. No. 3.1e-110;
Matches 445; Conservative 0; Mismatches 4; Indels 4; Gaps 4;

Qy 1 atcaagaacatagagttcgggcaatacttcatctcactccatccaccaccccaaatctta 60
Db 709 ATCAAGAAACATAGAGTT-GGGCAATATATCTTCCATCTACCCATCCCAAAATCTTA 651
Qy 61 ctctactcattctcatttaatttggaaatcatcagaagatgtgtcttgagta 120
Db 650 CTCTACTCATCTCATCTCAATTAATTTGGGAATCATCAGAAATGTTGCTTGAGTA 591
Qy 121 agagattaaagaataagctttttgacccctgccaacacccatgccagggtggtcac 180
Db 590 AGAGATTAAGAATAAGCTTTTGAACCCCTGCCAACACCCATGCCCAGGGTGTAC 531
Qy 181 cttcaatacaataacatccaggaagatgaattggcctttctgtatccgtaactgcc 240
Db 530 CTCCAATAACAATAGATGCCAGGAAGAGTAAGTTGCCCTTTCTGATGCCGTAATCTGCC 471
Qy 241 atcatcttcccatcttccagtcctcttccattgcaagtccaactctgggtctcaggat 300
Db 470 ATCATCTTCCCATCTTCCAGTCT-CTTTCCATTGCAAGTCAACAATCTGGGTCACGGAT 412
Qy 301 tataccgcttagtcgatcattgtcttcaactgtgccaactgccaggtggtcac 360
Db 411 TATACCCGCTTAGTCTCGATCATCTGCTTTCACCTTGTGCCACTGAGCTGGACCTTCGCAC 352
Qy 361 ctggaggaggtgctcttctgctcatcactgactcaccagaaacaaaggagcgtcct 420
Db 351 CT-GGAGGAGGTGCTCTTTGCTCTCATCAGTACTCCACAAGAAACAAAGGGCAGCTCCT 293
Qy 421 catcactggggtttcaccactttcagggttaag 453
Db 292 CATCACT-GGGCTTACCACCTTTTCAGGGTAAGG 261

RESULT 10
US-60-172-373-12767/c
; Sequence 12767, Application US/60172373
```

```

: ORGANISM: Homo sapiens
:
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: 208328.1
:
: NAME/KEY: unsure
: LOCATION: 1547-1579
:
: OTHER INFORMATION: a, t, c, g, or other
:
: US-60-209-009-485

```

Query Match	85.5%	Score 398.6:	DB 59:	Length 2697:
Best Local Similarity	98.2%	Pred. No. 5,2e-110:		
Matches 445:	Conservative 0:	Mismatches 4:	Indels 4:	Gaps
Qy	1	atcaagaaacatagagttcgggcaataactctaccatccaccaccaaacttta	60	
Db	709	ATCAAGAAACATAGAGTT-GGSCAATACTTCTATCTTACCATCCCAACCAATCTTA	651	
Qy	51	ctctactcatctcattctcatttaattttgggaaatcatcagaagatgtttcgttgagta	120	
Db	650	CTCTACTCATCTCATTTCTCAATAATTTTGGGAAATCATCAGAAGATGTGTTCTGTGAGTA	591	
Qy	121	agagattaaagaaaataaagctttttgacccctgccacccccatgccagagtggtgcac	180	
Db	590	AGAGATTAAAGAAATAATGAGCTTTTGACCCCTGCCAACCCCATGCCAGGGTGTGTAC	531	
Qy	181	cctccaatacaataaaccgacaggagaagtaagttgcccctttctgatgcggtaaactgcc	240	
Db	530	CCTCCAATACAATAAAGATGCGCAGSAGAGTAAGTTGCCCTTTCTGATGCGTAATCTGCC	471	
Qy	241	atcatctcccatctccagctctctttccatigcaagtcacaactctgggtctcaggagat	300	
Db	470	AICATCTTCCCATCTTCCAGTCT-CTTTCCATTGCAAGTCACAATCTGGTCTCAGGGAT	412	
Qy	301	tataccgcttagtctcgtcatctgttttcactttgtgcactgagctggagcccttcgac	360	
Db	411	TATACCCGCTCTTAGTCTCGATCATTTGCTTTTCTACTTTGCCCACCTGAGCTGGACCTTCGCAC	352	
Qy	361	ctgggagagagtgctcttttgctctcatcactgactccacaagaaacaaaggagcgtctct	420	
Db	351	CT-GGAGGAGGTGGCTCTTTGCCCTCATCACCCTGACTCTCAAGAAACAGGGCAGCTCTCT	293	
Qy	421	catcactggggcttcacccacttttcaggggtaag	453	
Db	263	CATCACT-EGGCTTCACACTTTCAGGGTAAGG	261	

```

RESULT 12
US-50-269-106-423/c
: Sequence 423, Application US/60209106
: GENERAL INFORMATION:
: APPLICANT: Kaser, Matthew R.
: APPLICANT: Bandman, Olga
: TITLE OF INVENTION: GENES EXPRESSED IN TNF ALPHA-TREATED VASCULAR ENDOTHELIUM AND
: TITLE OF INVENTION: MUSCLE CELLS
: FILE REFERENCE: PA-0031 P
: CURRENT APPLICATION NUMBER: US/60/209, 106
: CURRENT FILING DATE: 2000-06-02
: NUMBER OF SEQ ID NOS: 499
: SOFTWARE: PERL Program
: SEQ ID NO 423
: LENGTH: 2697
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: 208328.1
: FEATURE:
: NAME/KEY: unsure
: LOCATION: 1547-1579
: OTHER INFORMATION: a, t, c, g, or other
US-50-269-106-423

```

```

Qy   61 ctctactcatctcatttcattaattttgggaaatcatcaagaagatgttccgttgagta 120
Db   650 CTCCTACTCACTCATCTTCATTAATTTTGGGAAATCATCAGAAGATGTGTCTTGAGTA 591
Qy   121 agagattaaaagaaataaagctttttgaccccctgcacaaccccccatgccagggtggtcac 180
Db   590 AGAGATTAAAACAAATAAAGCTTTTTTGACCCCTGCACAACACCCCATGCCAGGGTGTCCAC 531
Qy   181 cctccaataacaatacatgcccagggaagtagtaagtgtccctttctgatcgctaactctgcc 240
Db   530 CTTCCAATAACAATAAGATGCCAGGAAGTAGTAAGTGTGCCCTTTCTGTATGCCGTAATCTGCC 471
Qy   241 atcatettcccatcttcacagtcctctttccattgcaagtcacaaatctgggtctcagggat 300
Db   470 ATCATCTTCCCATCTTCAGTCT-CTTTCCATTGCAAGTCACAATCTGGGTCTCAGGGAT 412
Qy   301 tataccgctcttagtcgatacttgctttcaccttgtgcactgagctggaccttcgcac 360
Db   411 TATACCCTGCTTAGTCTCGATCATTGCTTTCACTGTGCCACTGAGCTGAGCTTCGCAC 352
Qy   361 ctggaggagggtgctcttttgctcatcacctgactccacaagaacaaagggcagctcct 420
Db   351 CT-GGAGGAGGTGCCCTTTTGCTCTATCACCTGACTCCACAAGAAACAAGGCGAGCTCCT 293
Qy   421 catcactgggggtccaccacttcacaggggtaag 453
Db   292 CATCACT-GGGCTTACCACITTTACGGGTAAGG 261

RESULT 14
US-60-278-258-16025/c
: Sequence 16025, Application US/60278258
: GENERAL INFORMATION:
: APPLICANT: Morris, MacDonald
: APPLICANT: Lal, Preeti
: APPLICANT: Diep, Dinh
: TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
: TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
: TITLE OF INVENTION: Polymorphisms Identified Thereby
: FILE REFERENCE: GX-0010-1 P
: CURRENT APPLICATION NUMBER: US/60/278,258
: CURRENT FILING DATE: 2001-03-23
: NUMBER OF SEQ ID NOS: 17730
: SOFTWARE: PERL Program
: SEQ ID NO 16025
: LENGTH: 2697
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: 208328.1
: NAME/KEY: unsure
: LOCATION: 1547-1579
: OTHER INFORMATION: a, t, c, g, or other
US-60-278-258-16025

Query Match      85.5%; Score 398.6; DB 66; Length 2697;
Best Local Similarity 98.2%; Pred. No. 5,2e-110;
Matches 445; Conservative 0; Mismatches 4; Indels 4; Gaps 4;

Qy   1 atcaagaaacatagagttcgggcaatatactctatccatccaccaccaaacttta 60

```

61   gtctactcatctcatctcatttaattttgggaataatcatcaagaatgcttcgcttgagta   120  
          |||||  
650   CTTACTCATCTCTCATCTCTTAATTTTGGGAAATCATCAGAAGATGTGTTCTGTGAGTA   591  
          |||||

121   agagattaaagaataaagctttttgacccccctgcacaaaccccccatgccagggtgggtcac   180  
          |||||  
550   AGAGATTAAAGAANAATTAAGCTTTTGACCCCTTGCCAAACACCCCATGCCACGGGTGGGTAC   531  
          |||||





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 16:29:39 ; Search time 100.77 Seconds  
(without alignments)  
1135.906 Million cell updates/sec

Title: US-09-880-107-2492  
Perfect score: 466  
Sequence: 1 atcaagaacacatagattc.....gggtaagggtggatgtctt 466

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues  
Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43.4	9.3	7218	1	US-08-232-463-14
2	42.2	9.1	3840	1	US-08-462-092-1
3	42.2	9.1	3840	3	US-08-746-822-1
4	42.2	9.1	3840	3	US-09-094-350-1
5	42.2	9.1	3840	6	5510474-1
6	37.4	8.0	508	3	US-09-051-969A-1
7	36.6	7.9	2127	2	US-08-505-486-95
8	36.6	7.9	2127	3	US-08-801-028-95
9	36.6	7.9	2127	3	US-09-340-154-95
10	36.6	7.9	2127	5	PCT-US95-09338-95
11	36.6	7.9	2127	5	PCT-US95-09339-95
12	35.2	7.6	471	2	US-09-070-060-9
13	35.2	7.6	471	3	US-09-357-746-9
14	35	7.5	1228	2	US-08-505-486-92
15	35	7.5	1228	3	US-08-801-028-92
16	35	7.5	1228	3	US-09-340-154-92
17	35	7.5	1228	5	PCT-US95-09338-92
18	35	7.5	1228	5	PCT-US95-09339-92
19	34.6	7.4	581	2	US-08-834-306-22
20	34.6	7.4	581	4	US-08-993-674A-22
21	34.6	7.4	581	4	US-08-505-486-93
22	34.2	7.3	1154	3	US-08-801-028-93
23	34.2	7.3	1154	3	US-09-340-154-93
24	34.2	7.3	1154	5	PCT-US95-09338-93
25	34.2	7.3	1154	5	PCT-US95-09339-93
26	34.2	7.3	1154	2	US-08-505-486-96
27	34.2	7.3	2022	2	

C 28	34.2	7.3	2022	3	US-08-801-028-96	Sequence 96, Appl
C 29	34.2	7.3	2022	3	US-09-340-154-96	Sequence 96, Appl
C 30	34.2	7.3	2022	5	PCT-US95-09338-96	Sequence 96, Appl
C 31	34.2	7.3	2022	5	PCT-US95-09339-96	Sequence 96, Appl
C 32	33.8	7.3	231	1	US-08-450-834-1	Sequence 1, Appl
C 33	33.8	7.3	831	1	US-08-450-834-5	Sequence 5, Appl
C 34	33.6	7.2	487	3	US-09-051-969A-2	Sequence 2, Appl
C 35	32.4	7.0	4673	1	US-07-638-431-1	Sequence 1, Appl
C 36	32.4	7.0	4673	5	PCT-US92-00018-1	Sequence 22, Appl
C 37	31	6.7	707	2	US-08-726-306A-22	Sequence 6, Appl
C 38	30.8	6.6	418	3	US-09-141-000-6	Sequence 4, Appl
C 39	30.8	6.6	458	3	US-09-141-000-4	Sequence 24, Appl
C 40	30.8	6.6	1501	2	US-08-145-658D-24	Sequence 73, Appl
C 41	30.2	6.5	15766	4	US-09-338-907-73	Sequence 73, Appl
C 42	30.2	6.5	15766	4	US-09-218-207-73	Sequence 183, App
C 43	30.2	6.5	37950	4	US-09-338-907-183	Sequence 183, App
C 44	30.2	6.5	37950	4	US-09-218-207-183	Sequence 30, Appl
C 45	30	6.4	1622	1	US-08-216-276A-30	

ALIGNMENTS

RESULT 1  
: Sequence 14, Application US/08232463  
: Patent No. 5670367  
: GENERAL INFORMATION:  
: APPLICANT: DORNER, F.  
: APPLICANT: SCHEIFLINGER, F.  
: APPLICANT: FALKNER, F. G.  
: TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
: NUMBER OF SEQUENCES: 52  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Foley & Lardner  
: STREET: 1800 Diagonal Road, Suite 500  
: CITY: Alexandria  
: STATE: VA  
: COUNTRY: USA  
: ZIP: 22313-0299  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent in Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/232,463  
: FILING DATE:  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US/07/935,313  
: FILING DATE:  
: APPLICATION NUMBER: EP 91 114 300.6  
: FILING DATE: 26-AUG-1991  
: ATTORNEY/AGENT INFORMATION:  
: NAME: BENT, Stephen A.  
: REGISTRATION NUMBER: 29,768  
: REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (703)836-9300  
: TELEFAX: (703)683-4109  
: TELEX: 899149  
: INFORMATION FOR SEQ ID NO: 14:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 7218 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: IMMEDIATE SOURCE:  
: CLONE: pizgpt-Fls  
: US-08-232-463-14

Query Match	9.3%	Score 43.4	DB 1	Length 7218
Best Local Similarity	2.8%	Pred. No. 9.3e-05		
Matches	11	Conservative 216	Mismatches 162	Indels 0
Gaps	0			
14 agagtctggcgcaataactctatcctaccatccaccacaaatcttactctactc 73				
1053 AGGAGCTTCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1112				
74 attctcattaat1ttgggaaatcatcagaagatgtttcgttgagtaagagattaaaaga 133				
1113 YY 1172				
134 aataagcttttgacccctgccaaccccatgccagggtgcacctccaatacaat 193				
1173 YY 1232				
194 aacatgcaggagaagtaagtgcctttctgtagcgcgtaatctgcacatcattctcccat 253				
1233 YY 1292				
254 ctccagctcctcttcacatgcaagtcacaatctgggtctcagggaattatcccgctcta 313				
1293 YY 1352				
314 gctcgcattctttcaccttgccacatgagctgcaccttcgcacctggaggagggtg 373				
1353 YY 1412				
374 cctctttgcctcattcaactgactccacaa 402				
1413 YYGTACCAA 1441				

```

1  RESULT      2
2  US-08-462-092-1/c
3  ; Sequence 1, Application US/08462092
4  ; Patent No. 5614399
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Peter H. Quail
7  ; APPLICANT: Alan H. Christensen
8  ; APPLICANT: Howard P. Hershey
9  ; APPLICANT: Robert A. Sharrack
10 ; APPLICANT: Thomas D. Sullivan
11 ; TITLE OF INVENTION: PLANT UBIQUITIN PROMOTER SYSTEM
12 ; NUMBER OF SEQUENCES: 2
13 ; CORRESPONDENCE ADDRESS:
14 ; ADDRESSEE: Jeff Lloyd
15 ; STREET: 2421 N.W. 41st Street, Suite A-1
16 ; CITY: Gainesville
17 ; STATE: FL
18 ; COUNTRY: USA
19 ; ZIP: 32606
20 ; COMPUTER READABLE FORM:
21 ; MEDIUM TYPE: Floppy disk
22 ; COMPUTER: IBM PC compatible
23 ; OPERATING SYSTEM: PC-DOS/MS-DOS
24 ; SOFTWARE: Patent In Release #1.0, Version #1.25
25 ; CURRENT APPLICATION DATA:
26 ; APPLICATION NUMBER: US/08/462,092
27 ; FILING DATE: 5-JUNE-1995
28 ; CLASSIFICATION: 435
29 ; ATTORNEY/AGENT INFORMATION:
30 ; NAME: Lloyd, Jeff
31 ; REGISTRATION NUMBER: 35,589
32 ; REFERENCE/DOCKET NUMBER: 08/462,092
33 ; TELECOMMUNICATION INFORMATION:
34 ; TELEPHONE: 904-375-8100
35 ; TELEFAX: 904-372-5800
36 ; INFORMATION FOR SEQ ID NO: 1:
37 ; SEQUENCE CHARACTERISTICS:
38 ; LENGTH: 3840 base pairs
39 ; TYPE: nucleic acid
40 ; STRANDEDNESS: single

```

```

: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1993..3591
: US-08-462-092-1

Query Match          9.1%; Score 42.2; DB 1; Length 3840;
Best Local Similarity 55.1%; Pred. No. 0.00017;
Matches 103; Conservative 0; Mismatches 83; Indels 1; Gaps

QY 200 ccaggagagtaagtgccttcttgatgccgtaatactgccaatacttcccatcttcca 259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3568 CAAAGTGGAGGTGCTCTCCTTTGATGTTGTAGTCTGCTAGGTGCGGCATCTCTCCA 3509

QY 260 gtctctttccattgcaatcacaatctgggtctcaggattataccgctcttagtctcg 319
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3559 G-CGTGTTGCCGCGAAGATCAGACGCTGCTGGTCCGGGGGATGCCCTCTCTGCTCTGG 3450

QY 320 atcattgtcttactgtgccctagctgagcttgcaccttcgcacctgggagaggtgcctctt 379
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3449 ATCTTCGCTTCACATTGTCGATGTTGCCGAGCTCTCAACCTCCAGGGTTATGGTITTT 3390

QY 380 tgcctca 386
    |||
Db 3389 CCAGTCA 3383

RESULT 3
US-08-746-822-1/c
: Sequence 1, Application US/08746822
: Patent No. 6020190
: GENERAL INFORMATION:
: APPLICANT: Peter H. Quail
: APPLICANT: Alan H. Christensen
: APPLICANT: Howard P. Hershey
: APPLICANT: Robert A. Sharrock
: APPLICANT: Thomas D. Sullivan
: TITLE OF INVENTION: PLANT UBQUITIN PROMOTER SYSTEM
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Jeff Lloyd
: STREET: 2421 N.W. 41st Street, Suite A-1
: CITY: Gainesville
: STATE: FL
: COUNTRY: USA
: ZIP: 32606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/746,822
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/462,092
: FILING DATE: 5-JUNE-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Lloyd, Jeff
: REGISTRATION NUMBER: 35,589
: REFERENCE/DOCKET NUMBER: 08/462,092
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 904-375-8100
: TELEFAX: 904-372-5800
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3840 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single

```

TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1993..3591  
US-08-746-822-1

Query Match 9.1% Score 42.2; DB 3; Length 3840;

Best Local Similarity 55.1% Pred. No. 0.00017;  
Matches 103; Conservative 0; Mismatches 83; Indels 1; Gaps 1;

Oy 200 ccaggaagagtaagtgcctttctgagtcgcgtaataatcgccatcatcttcccatctcca 259  
Db 3568 CAAGGTGGAGGGTCTCTCTTTTGGATGTTGTAGTCTGCTAGGCTGGCGGCATCTCCA 3509  
Oy 260 gtctcttccattgcaatgcaatctggtctcaaggattataccgctttagctcg 319  
Db 3508 G-CTGTTGCGCGGAAGATCAGACGCTGCTGTCGGGGGATGCCCTCTTGTCTTGG 3450  
Oy 320 atcattgcttccattgctccactgagctggaccttcgcacctcgaggaggtgcctctt 379  
Db 3449 ATCTCGCTTCACATTTGTCGATGTTGTCGAGCTCTCAACCTCCAGGGTTATGGTTT 3390  
Oy 380 tgcctca 386  
Db 3389 CCAGTCA 3383

## RESULT 4

US-09-094-350-1/c  
Sequence 1, Application US/09094350  
Patent No. 6054574  
GENERAL INFORMATION:  
APPLICANT: Peter H. Quail  
APPLICANT: Alan H. Christensen  
APPLICANT: Howard P. Hershey  
APPLICANT: Robert A. Sharruck  
APPLICANT: Thomas D. Sullivan  
TITLE OF INVENTION: PLANT UBIQUITIN PROMOTER SYSTEM  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jeff Lloyd  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: USA  
ZIP: 32606  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION NUMBER: US/09/094,350  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/746,822  
FILING DATE:  
APPLICATION NUMBER: US/08/462,092  
FILING DATE: 5-JUNE-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Lloyd, Jeff  
REGISTRATION NUMBER: 35,589  
REFERENCE/DOCKET NUMBER: 08/462,092  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 904-375-8100  
TELEFAX: 904-372-5800  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3840 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1993..3591  
US-09-094-350-1

Query Match 9.1% Score 42.2; DB 3; Length 3840;

Best Local Similarity 55.1% Pred. No. 0.00017;  
Matches 103; Conservative 0; Mismatches 83; Indels 1; Gaps 1;

Oy 200 ccaggaagagtaagtgcctttctgagtcgcgtaataatcgccatcatcttcccatctcca 259  
Db 3568 CAAGGTGGAGGGTCTCTCTTTTGGATGTTGTAGTCTGCTAGGCTGGCGGCATCTCCA 3509  
Oy 260 gtctcttccattgcaatgcaatctggtctcaaggattataccgctttagctcg 319  
Db 3508 G-CTGTTGCGCGGAAGATCAGACGCTGCTGTCGGGGGATGCCCTCTTGTCTTGG 3450  
Oy 320 atcattgcttccattgctccactgagctggaccttcgcacctcgaggaggtgcctctt 379  
Db 3449 ATCTCGCTTCACATTTGTCGATGTTGTCGAGCTCTCAACCTCCAGGGTTATGGTTT 3390  
Oy 380 tgcctca 386  
Db 3389 CCAGTCA 3383

## RESULT 5

5510474-1/c  
Patent No. 5510474  
APPLICANT: QUAIL, PETER H.; CHRISTENSEN, ALAN H.; HERSHEY,  
HOWARD P.; SHARRUCK, ROBERT A.; SULLIVAN, THOMAS D.  
TITLE OF INVENTION: PLANT UBIQUITIN PROMOTER SYSTEM  
NUMBER OF SEQUENCES: 12  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/296,268  
FILING DATE: 25-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 191,134  
FILING DATE: 03-FEB-1994  
APPLICATION NUMBER: 76,363  
FILING DATE: 11-JUN-1993  
APPLICATION NUMBER: 670,496  
FILING DATE: 15-MAR-1991  
APPLICATION NUMBER: 194,824  
FILING DATE: 17-MAY-1988  
SEQ ID NO: 1:  
LENGTH: 3840  
5510474-1

Query Match 9.1% Score 42.2; DB 6; Length 3840;

Best Local Similarity 55.1% Pred. No. 0.00017;  
Matches 103; Conservative 0; Mismatches 83; Indels 1; Gaps 1;

Oy 200 ccaggaagagtaagtgcctttctgagtcgcgtaataatcgccatcatcttcccatctcca 259  
Db 3568 CAAGGTGGAGGGTCTCTCTTTTGGATGTTGTAGTCTGCTAGGCTGGCGGCATCTCCA 3509  
Oy 260 gtctcttccattgcaatgcaatctggtctcaaggattataccgctttagctcg 319  
Db 3508 G-CTGTTGCGCGGAAGATCAGACGCTGCTGTCGGGGGATGCCCTCTTGTCTTGG 3450  
Oy 320 atcattgcttccattgctccactgagctggaccttcgcacctcgaggaggtgcctctt 379  
Db 3449 ATCTCGCTTCACATTTGTCGATGTTGTCGAGCTCTCAACCTCCAGGGTTATGGTTT 3390  
Oy 380 tgcctca 386



APPLICANT: Jesse M. Jaynes  
TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE  
TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND  
TITLE OF INVENTION: METHODS OF MAKING AND USING SAME  
NUMBER OF SEQUENCES: 98  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ  
STREET: 555 Thirteenth Street N.W.  
CITY: Washington  
STATE: D. C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE  
OPERATING SYSTEM: IBM COMPATIBLE  
SOFTWARE: WordPerfect 5.1+  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/340,154  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/505,486  
FILING DATE: 21-JUL-1995  
APPLICATION NUMBER: U.S. 08/279,472  
FILING DATE: 22-JUL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: WALKER, BARBARA W.  
REGISTRATION NUMBER: 35,400  
REFERENCE/DOCKET NUMBER: 2093-117A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)783-6040  
TELEFAX: (202)783-6031  
INFORMATION FOR SEQ ID NO: 95:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2127  
TYPE: NUCLEIC ACID  
STRANDEDNESS: DOUBLE STRANDED  
TOPOLOGY: LINEAR  
MOLECULE TYPE: GENOMIC DNA AND OTHER DNA  
DESCRIPTION: GENOMIC DNA AND OTHER DNA  
US-09-340-154-95  
Query Match 7.9% Score 36.6; DB 3; Length 2127;  
Best Local Similarity 50.2%; Pred. No. 0.011;  
Matches 116; Conservative 0; Mismatches 114; Indels 1; Gaps 1;  
QY 142 ttgtgacccctgcccaacacccatgcccaggggtggtcaccctccaataacaataacatgcc 201  
DB 2054 TTTTCAACCTACGACCCACTCTTTTAAACAGCGGATCCACCTCCACGTAGAGGAGCACC 1995  
QY 202 aggaagagtaagtgccctttctgacggtggtcaccctccaataacaataacatgcc 261  
DB 1994 AAATGGAGGGTAGACTCTCTTGGATGTTGTAATCAGCTAGAGTAGCTCCGCTCTCCA-A 1936  
QY 262 ctctttccattgcaagtcacaaatctgggtctcagggattataccgctcttagtctcgat 321  
DB 1935 CTGCTTTCGGCGAAGATAAGCCCTTTCGTATCCGGGGAAATTCCTTATCTCTGGAT 1876  
QY 322 cattgtcttcaacttgccactgagctggacccctgcacactggaggaggt 372  
DB 1875 CTTAGCCCTTAACGTTGTCGATGATCATCAGAACTTTCACCTCTAGGGTGTAT 1825  
RESULT 10  
PCT-US95-09338-95/c  
Sequence 95, Application PC/TUS9509338  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE  
TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND  
TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME

ADDRESSEE: STEVEN J. HULTQUIST  
ADDRESSEE: INTELLECTUAL PROPERTY/TECHNOLOGY LAW  
STREET: 200 PARK DRIVE, SUITE 210  
STREET: P.O. BOX 14329  
CITY: RESEARCH TRIANGLE PARK  
STATE: NORTH CAROLINA  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE  
OPERATING SYSTEM: APPLE MACINTOSH  
SOFTWARE: M.S. WORD 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/801,028  
FILING DATE: 19-FEB-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/279,472  
FILING DATE: JULY 22, 1994  
APPLICATION NUMBER: 08/225,476  
FILING DATE: 04-20-94  
APPLICATION NUMBER: 08/225,476  
FILING DATE: 04-08-94  
APPLICATION NUMBER: 08/039,620  
FILING DATE: 06-04-93  
APPLICATION NUMBER: 08/148,491  
FILING DATE: 11-08-93  
APPLICATION NUMBER: 08/148,889  
FILING DATE: 11-08-93  
ATTORNEY/AGENT INFORMATION:  
NAME: WASSERMAN, FRAN S.  
REGISTRATION NUMBER: 34,273  
REFERENCE/DOCKET NUMBER: 4013-104  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)990-9531  
TELEFAX: (919)990-9532  
INFORMATION FOR SEQ ID NO: 95:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2127  
TYPE: NUCLEIC ACID  
STRANDEDNESS: DOUBLE STRANDED  
TOPOLOGY: LINEAR  
MOLECULE TYPE: GENOMIC DNA AND OTHER DNA  
DESCRIPTION: GENOMIC DNA AND OTHER DNA  
US-08-801-028-95  
Query Match 7.9% Score 36.6; DB 3; Length 2127;  
Best Local Similarity 50.2%; Pred. No. 0.011;  
Matches 116; Conservative 0; Mismatches 114; Indels 1; Gaps 1;  
QY 142 ttgtgacccctgcccaacacccatgcccaggggtggtcaccctccaataacaataacatgcc 201  
DB 2054 TTTTCAACCTACGACCCACTCTTTTAAACAGCGGATCCACCTCCACGTAGAGGAGCACC 1995  
QY 202 aggaagagtaagtgccctttctgacggtggtcaccctccaataacaataacatgcc 261  
DB 1994 AAATGGAGGGTAGACTCTCTTGGATGTTGTAATCAGCTAGAGTAGCTCCGCTCTCCA-A 1936  
QY 262 ctctttccattgcaagtcacaaatctgggtctcagggattataccgctcttagtctcgat 321  
DB 1935 CTGCTTTCGGCGAAGATAAGCCCTTTCGTATCCGGGGAAATTCCTTATCTCTGGAT 1876  
QY 322 cattgtcttcaacttgccactgagctggacccctgcacactggaggaggt 372  
DB 1875 CTTAGCCCTTAACGTTGTCGATGATCATCAGAACTTTCACCTCTAGGGTGTAT 1825  
RESULT 9  
US-09-340-154-95/c  
Sequence 95, Application US/09340154  
Patent No. 6084156  
GENERAL INFORMATION:

```
/ NUMBER OF SEQUENCES: 98
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: WORDPERFECT 5.1+
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/09338
/ PRIOR APPLICATION DATA: 08/279,472
/ APPLICATION NUMBER: 08/279,472
/ FILING DATE: 22-JUL-1994
/ INFORMATION FOR SEQ ID NO: 95:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2127
/ TYPE: NUCLEIC ACID
/ STRANDEDNESS: DOUBLE STRANDED
/ TOPOLOGY: LINEAR
/ MOLECULE TYPE:
/ DESCRIPTION: GENOMIC DNA AND OTHER DNA
PCT-US95-09338-95

Query Match 7.9%; Score 36.6; DB 5; Length 2127;
Best Local Similarity 50.2%; Pred. No. 0.011;
Matches 116; Conservative 0; Mismatches 114; Indels 1; Gaps 1;

QY 142 tttagaccctgtccaacacccatgccagggtggtcaccctcccaataacaatgcc 201
Db 2054 TTTTCAACCTAGCAGCCACTCTTTTAAACAGCGGATCCACCTCCAGTAGAGGAGCACC 1995
QY 202 aggaagagtaagtgtccctttctgtatgccgttaattcgtcattcctccatcttccagt 261
Db 1994 AAATGGAGGTAGACTCCTTCTGGATGTTGTAATCAGCTAGAGTACGTCCTCTCCA-A 1936
QY 262 ctcttctcattgtgcaagtcacaatctctgggtctcaggagattataccctcttctcagat 321
Db 1935 CTGCTTTCCGGCGAAGATAAGCCTTTCTGATCGCGGGAATTCCTTCTTATCTCTGGAT 1876
QY 322 cattgttctcattgtgcaagtcacaatctctgggtctcaggagattataccctcttctcagat 372
Db 1875 CTTAGCCTTAACGTTGTCGATTGATCAGAACTTTCCACCTCTAGGGTGAT 1825

RESULT 11
PCT-US95-09339-95/c
; Sequence 95, Application PC/TUS9509339
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
; TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
; NUMBER OF SEQUENCES: 98
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09339
; FILING DATE: 21-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/279,472
; FILING DATE: 22-JUL-1994
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2127
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE STRANDED
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: GENOMIC DNA AND OTHER DNA
```

```
PCI-US95-09339-95

Query Match 7.9%; Score 36.6; DB 5; Length 2127;
Best Local Similarity 50.2%; Pred. No. 0.011;
Matches 116; Conservative 0; Mismatches 114; Indels 1; Gaps 1;

QY 142 tttagaccctgtccaacacccatgccagggtggtcaccctcccaataacaatgcc 201
Db 2054 TTTTCAACCTAGCAGCCACTCTTTTAAACAGCGGATCCACCTCCAGTAGAGGAGCACC 1995
QY 202 aggaagagtaagtgtccctttctgtatgccgttaattcgtcattcctccatcttccagt 261
Db 1994 AAATGGAGGTAGACTCCTTCTGGATGTTGTAATCAGCTAGAGTACGTCCTCTCCA-A 1936
QY 262 ctcttctcattgtgcaagtcacaatctctgggtctcaggagattataccctcttctcagat 321
Db 1935 CTGCTTTCCGGCGAAGATAAGCCTTTCTGATCGCGGGAATTCCTTCTTATCTCTGGAT 1876
QY 322 cattgttctcattgtgcaagtcacaatctctgggtctcaggagattataccctcttctcagat 372
Db 1875 CTTAGCCTTAACGTTGTCGATTGATCAGAACTTTCCACCTCTAGGGTGAT 1825

RESULT 12
US-09-070-960-9/c
; Sequence 9, Application US/09070060
; GENERAL INFORMATION:
; APPLICANT: Hustad, Carolyn M.
; APPLICANT: Ghildyal, Namit
; TITLE OF INVENTION: Human E3 Ubiquitin Protein
; TITLE OF INVENTION: Ligase
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA Pharmaceuticals, Inc.
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: USA
; ZIP: 19850-5437
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,060
; FILING DATE: 30-APR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/073,839
; FILING DATE: 05-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Higgins, Patrick H
; REGISTRATION NUMBER: 39,709
; REFERENCE/DOCKET NUMBER: PHM.70312
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302.886.4889
; TELEFAX: 302.886.8221
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 471 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
US-09-070-360-9

Query Match 7.6%; Score 35.2; DB 2; Length 471;
Best Local Similarity 55.8%; Pred. No. 0.013;
```

```

RESULT 14
US-08-505-486-92/c
: Sequence 92, Application US/08505486
: Patent No. 5955573
: GENERAL INFORMATION:
: APPLICANT: Jesse M. Jaynes
: TITLE OF INVENTION: UBQUITIN-LYTIC PEPTIDE FUSION GENE
: TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
: TITLE OF INVENTION: METHODS OF MAKING AND USING SAME
: NUMBER OF SEQUENCES: 98
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
: STREET: 555 Thirteenth Street N.W.
: CITY: Washington
: STATE: D. C.
: COUNTRY: USA
: ZIP: 20004
: COMPUTER READABLE FORM:

```

```

: GENERAL INFORMATION:
: APPLICANT: JOAN GARBARINO
: APPLICANT: JESSE M. JAYNES
: APPLICANT: WILLIAM BELNAP
: TITLE OF INVENTION: UBQUITIN-LYTIC PEPTIDE FUSION GENE CONSTRUCTS, PROTEIN PRO
: NUMBER OF SEQUENCES: 98
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: STEVEN J. HULTQUIST
: ADDRESSEE: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
: STREET: 200 PARK DRIVE, SUITE 210
: STREET: P.O. BOX 14329
: CITY: RESEARCH TRIANGLE PARK
: STATE: NORTH CAROLINA
: COUNTRY: USA
: ZIP: 27709
: COMPUTER READABLE FORM:
: MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
: COMPUTER: APPLE MACINTOSH

```



```

; OPERATING SYSTEM: MACINTOSH
; SOFTWARE: M.S. WORD 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801.028
; FILING DATE: 19-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/279,472
; FILING DATE: JULY 22, 1994
; APPLICATION NUMBER: 08/225,476
; FILING DATE: 04-20-94
; APPLICATION NUMBER: 08/225,476
; FILING DATE: 04-08-94
; APPLICATION NUMBER: 08/039,620
; FILING DATE: 06-04-93
; APPLICATION NUMBER: 08/148,491
; FILING DATE: 11-08-93
; APPLICATION NUMBER: 08/148,889
; FILING DATE: 11-08-93
; ATTORNEY/AGENT INFORMATION:
; NAME: WASSERMAN, FRAN S.
; REGISTRATION NUMBER: 34,273
; REFERENCE/DOCKET NUMBER: 4013-104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)990-9531
; TELEFAX: (919)990-9532
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1228
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE STRANDED
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: GENOMIC DNA AND OTHER NUCLEIC ACID
; US-08-801-028-92

Query Match 7.5%; Score 35; DB 3; Length 1228;
Best Local Similarity 49.8%; Pred. No. 0.028;
Matches 115; Conservative 0; Mismatches 115; Indels 1; Gaps 1;

QY 142 tttgacctgcccaacaccccatgccagggtggtcacccctccaataacaataacatgcc 201
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1186 TTTTCAACCTACGACCCACTCTTTTAACAGCGGATCCACCACCGGACGAGCAGCAG 1127

QY 202 aggaagagtaagtgcctttctgatcgccgaatcgccatcatcttccatctccagt 261
    || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1126 AGATGGAGAGTTGACTCCTCTCTGGAITTTAGTCGGCAAGAGTACGACCACTCTCAAG- 1056

QY 262 ctctttccattgcaagtcaaatctgggtctcagggattataacccgtcttagtctcgat 321
    || |||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 1067 CTGCTTTCGGCGAAATCAAAACGCTGCTGCTGTTGGGGGAATCCCTTCCTTGTCIGGAT 1008

QY 322 catgtcttcactgtgcactgagctgacaccttgcacctgggagaggt 372
    || | |||| | | | | | | | | | | | | | | | | | | | | | |
Db 1007 CTGCTTTTGACATGTGATGTTGTCGGAGAGACTCAACCTCTAGGGTGAT 957
```

Search completed: July 29, 2002, 17:54:40  
Job time: 5101 sec